Box Tomography: first application to the imaging of upper-mantle shear velocity and radial anisotropy structure beneath the North American continent

Pierre Clouzet, Yder J. Masson, Barbara Romanowicz

Geophysical Journal International

Geophys. J. Int. (2018) **213,** 1849–1875
Advance Access publication 2018 February 27
GJI Seismology

doi: 10.1093/gji/ggy078

Box Tomography: first application to the imaging of upper-mantle shear velocity and radial anisotropy structure beneath the North American continent

P. Clouzet,[1] Y. Masson[1,*] and B. Romanowicz[1,2,3]

[1] *Departement de Sismologie, Institut de Physique du Globe de Paris,* 1 *rue Jussieu, F-*75005 *Paris, France. E-mail: clouzet.pierre@gmail.com*
[2] *Collège de France,* 11 *Place Marcelin Berthelot, F-*75231 *Paris, France*
[3] *Berkeley Seimological Laboratory, UC Berkeley, McCone Hall,* 215 *Haviland Path* #4760*, Berkeley, CA* 94720*, United States*

Accepted 2018 February 23. Received 2018 February 5; in original form 2017 June 26

SUMMARY

The EarthScope Transpotable Array (TA) deployment provides dense array coverage throughout the continental United States and with it, the opportunity for high-resolution 3-D seismic velocity imaging of the stable part of the North American (NA) upper mantle. Building upon our previous long-period waveform tomographic modeling, we present a higher resolution 3-D isotropic and radially anisotropic shear wave velocity model of the NA lithosphere and asthenosphere. The model is constructed using a combination of teleseismic and regional waveforms down to 40 s period and wavefield computations are performed using the spectral element method both for regional and teleseismic data. Our study is the first tomographic application of 'Box Tomography', which allows us to include teleseismic events in our inversion, while computing the teleseismic wavefield only once, thus significantly reducing the numerical computational cost of several iterations of the regional inversion. We confirm the presence of high-velocity roots beneath the Archean part of the continent, reaching 200–250 km in some areas, however the thickness of these roots is not everywhere correlated to the crustal age of the corresponding cratonic province. In particular, the lithosphere is thick (∼250 km) in the western part of the Superior craton, while it is much thinner (∼150 km) in its eastern part. This may be related to a thermomechanical erosion of the cratonic root due to the passage of the NA plate over the Great Meteor hotspot during the opening of the Atlantic ocean 200–110 Ma. Below the lithosphere, an upper-mantle low-velocity zone (LVZ) is present everywhere under the NA continent, even under the thickest parts of the craton, although it is less developed there. The depth of the minimum in shear velocity has strong lateral variations, whereas the bottom of the LVZ is everywhere relatively flat around 270–300 km depth, with minor undulations of maximum 30 km that show upwarping under the thickest lithosphere and downwarping under tectonic regions, likely reflecting residual temperature anomalies. The radial anisotropy structure is less well resolved, but shows distinct signatures in highly deformed regions of the lithosphere.

Key words: Composition and structure of the mantle; Seismic tomography.

1 INTRODUCTION

The North American continent (NA) consists of an ancient cratonic core, formed during the Archean, that has not been subjected to orogeny for at least 1.8 Ga, bordered to the southeast and east, by progressively younger, stable, Proterozoic provinces (Hoffman 1988). In the west, the Rocky Mountain Front (RMF) separates the Proterozoic and Archean shield from younger, tectonically active provinces (Fig. 1). This relatively regular geological configuration, combined with good seismic station coverage, makes NA an ideal target for seismic tomography, in order to investigate the relationship of lithospheric and asthenospheric structures to the geological features observed at the surface.

Differences in the seismic velocity structure between the craton and the western United States (US) extending down to at least 200 km were established already in the 1960s and 1970s from the first studies of teleseismic *P*- and *S*-wave traveltime anomalies (e.g. Cleary & Hales 1966; Herrin & Taggart 1968; Poupinet

* Now at: INRIA Bordeaux Sud-Ouest, Magique-3D Team, Pau, France; Université de Pau et des pays de l'Adour, France.

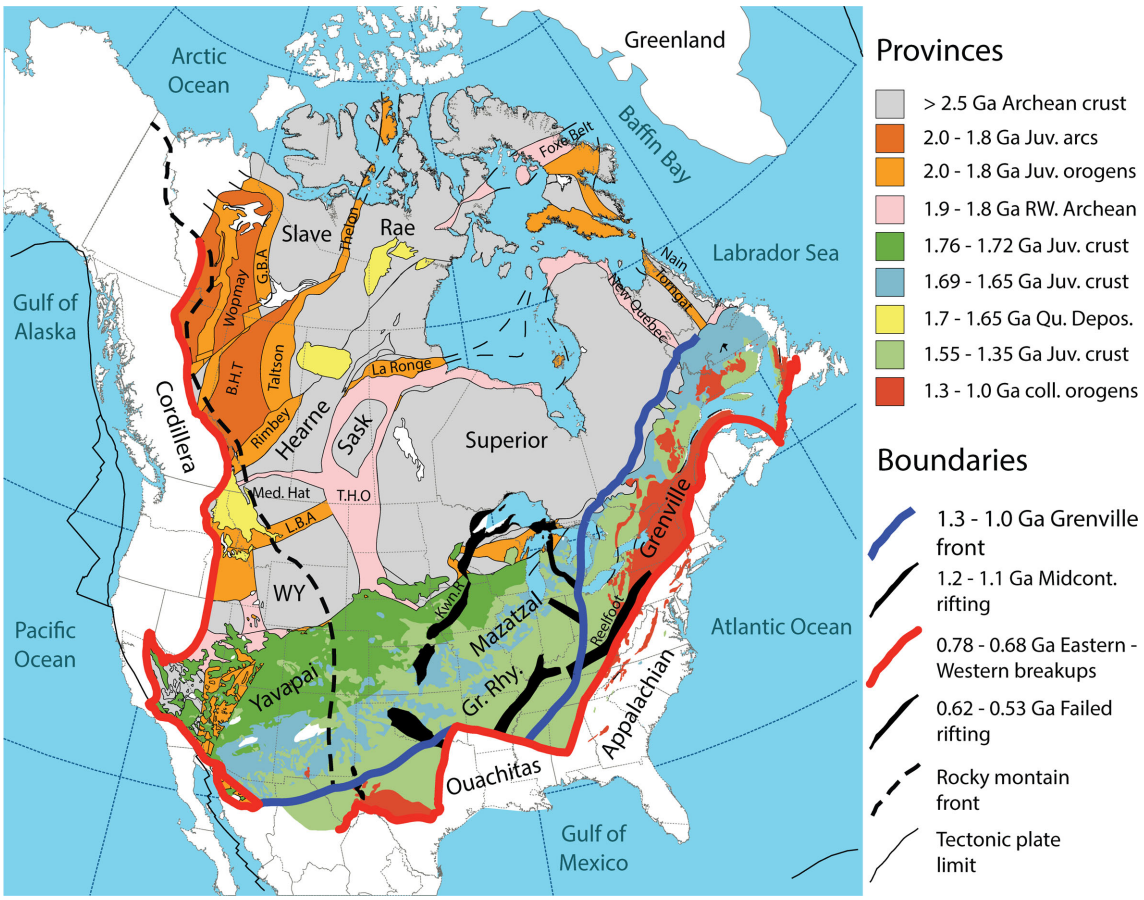

Figure 1. Large-scale Precambrian architecture of North America (modified from Whitmeyer & Karlstrom (2007). Colours indicate ages of different regions, as shown in the legend. The stable cratonic core is defined by the Rocky Mountain Front to the west (dashed dark grey line), defined by the surface trace of deformation), and the Appalachian front to the east (solid light red line). The solid blue line demarcates the westward extent of the Grenville/Llano Front. The solid red line along the Cordillera province demarcates the breakup of Rodinia. Solid green lines indicate tectonic plate boundaries. Abbreviations are as follows: RW, re-worked; GBA, Great Bear Arc; BHT, Buffalo Head Terrane; Taltson, Talston Arc; Rimbey, Rimbey Arc; Tornagat, Torngat Arc; L.B.A, Little Belt Arc; LRA; THO, Trans-Hudson Orogen; RGR, Rio Grande Rift; Reelfoot, Reelfoor Rift; Kwn.R, Kewanowan Rift; Sask, Sask Craton; WY, Wyoming Province; Med. Hat., Medicine Hat Block; Grt. Rhy, Granite-Rhyolite Province; Grenville, Grenville Province.

1979), the first tomographic images based on *P* and *S* traveltimes (Romanowicz 1979) and long-period surface waveforms (Woodhouse & Dziewonski 1984), as well as contrasted velocity–depth profiles obtained from the modeling of shear body waveforms at relatively short distances (e.g. Grand & Helmberger 1984). More recent body wave tomographic models have added considerable details to these images (Sigloch 2011; Burdick *et al.* 2014; Porritt *et al.* 2014, 2015).

Lateral variations in absolute velocities can be inferred from shear wave tomography, at scales from global (e.g. Su *et al.* 1994; Mégnin & Romanowicz 2000; Shapiro & Ritzwoller 2002; Panning & Romanowicz 2006; Kustowski *et al.* 2008; Ritsema *et al.* 2010; Lekić & Romanowicz 2011; Debayle & Ricard 2012; Schaeffer & Lebedev 2013) to continental (van der Lee & Frederiksen 2005; Yuan *et al.* 2011, 2014; Schaeffer & Lebedev 2014). These studies confirm the presence of strong lateral variations in the thickness of the lithosphere across the RMF, with lithospheric roots extending down to 200–250 km under the craton, thinning abruptly to the west to less than 80 km. This is also found from azimuthal anisotropy tomography, where a change in the anisotropy fast axis direction is observed across the lithosphere–asthenosphere boundary (Marone & Romanowicz 2007; Yuan & Romanowicz 2010).

While the mechanism of craton formation is still widely debated (e.g. King 2005; Lee *et al.* 2011), the presence of finer scale lateral and depth variations (Sigloch 2011; Burdick *et al.* 2017) in seismic structure suggest a complex history. Moreover, studies of azimuthal anisotropy have shown the presence of laterally varying layering within the cratonic lithosphere (e.g. Levin *et al.* 1999; Deschamps *et al.* 2008; Yuan & Romanowicz 2010), which may indicate different modes and/or times of formation of the top ∼100 km of this lithosphere. Fine-scale studies of converted and reflected phases indicate the presence of a sharp mid-lithospheric discontinuity (MLD) within the craton, marking the top of a mid-lithospheric low-velocity zone (LVZ, Thybo & Perchuć 1997; Bostock 1998; Abt *et al.* 2010; Fischer *et al.* 2010; Rader *et al.* 2015; Ford *et al.* 2016) and even possibly several MLDs (Calò *et al.* 2016). In addition, there is evidence from radially anisotropic shear wave tomography for separation between blocks of different crustal ages extending to depths of at least 150 km (e.g. Yuan *et al.* 2014).

Much of this finer scale structure has been resolved owing to the availability of data from the dense broad-band USArray TA deployment. With the completion of the coverage of the conterminous US, as well as availability of data from Canada and Greenland, it is now possible to further refine shear wave tomographic images of NA, and in particular of the stable Proterozoic and Archean provinces,

to try and improve our understanding of the different stages of formation of the continent. It is also an unprecedented opportunity to experiment with improved waveform-based tomographic techniques.

In a previous study, we presented a radially anisotropic shear velocity model of the NA upper mantle based on a combination of long-period teleseismic and regional waveform data (Yuan *et al.* 2014). The *regional* waveform data (down to 40 s period) were, for the first time for the NA continent, compared to 3-D synthetics computed using RegSEM (Cupillard *et al.* 2012), a spectral element method (SEM) code suitable for continental-scale wavefield computations. In contrast, due to the very heavy computations that would have been necessary to compute the predicted teleseismic wavefield numerically at each iteration of the inversion, the teleseismic wavefield was computed (down to 60 s period) using Non-Linear Assymptotic Coupling Theory (NACT, Li & Romanowicz 1995), a methodology based on normal-mode perturbation theory, that is computationally more efficient (albeit approximate), and has been used in the development of several generations of global and continental-scale shear velocity models based on time-domain waveform inversion (Li & Romanowicz 1996; Mégnin & Romanowicz 2000; Gung *et al.* 2003; Panning & Romanowicz 2006; Yuan *et al.* 2011).

The resulting 2014 radially anisotropic NA upper-mantle model presents some interesting features, in particular a correlation of radial anisotropy structure with lithospheric blocks corresponding to different orogenies in the eastern US and continental shelf. While more rigorous than inversions practiced by most groups and based on the path-average surface wave approximation (PAVA, Woodhouse & Dziewonski 1984), this methodology nevertheless presents some inconsistencies from the theoretical point of view, since the predicted wavefield through the target model space is computed with different theories for teleseismic (NACT) and regional (SEM) distance data, respectively.

To more consistently include teleseismic data into our regional tomographic inversions, we developed a general framework called 'Box Tomography' (see Masson *et al.* 2014; Masson & Romanowicz 2017a,b), that allows us to consistently model and invert both teleseismic (i.e. associated with sources or receivers outside of the imaged region) and regional (i.e. associated with sources and receivers within the imaged region) waveform data using accurate numerical methods such as SEM. In Box Tomography, prior to the inversion, the seismic wavefield generated by teleseismic sources is first modeled numerically at the global scale for a given reference 3-D model and is recorded at the surface of the region to be imaged. This reference wavefield is then used to construct virtual sources lying at the boundary of the regional modeling domain that are designed to reproduce the original wavefield within the target region, as illustrated in Figs 5 and 6. Once the teleseismic sources have been replaced by virtual sources within the regional modeling domain, the tomographic inversion can be performed efficiently (i.e. using regional modeling only) as the teleseismic data are accounted for seamlessly thanks to the virtual sources. Related concepts have been proposed to account for teleseismic data in regional imaging (e.g. Monteiller *et al.* 2015; Wang *et al.* 2016). However, in these studies, the global modeling of the reference wavefield is performed using faster, approximate methods, that do not account accurately for the effects induced on teleseismic waveforms by the 3-D structure of the Earth outside the regional modeling domain. In this paper, we present the first application of Box Tomography to continental-scale waveform tomography.

2 METHODOLOGY

Many studies have inverted for continental-scale structure, in particular in NA, using fundamental mode and, in some cases, overtone surface wave dispersion data or waveforms observed at teleseismic distances. The usual practice is to first consider the best possible global-scale tomographic shear wave velocity model, compute predictions of observables in this model using an approximate theory, generally the surface wave path-average approximation (e.g. Nettles & Dziewoński 2008; Bedle & van der Lee 2009; Schaeffer & Lebedev 2014), or, in our group, NACT (Marone *et al.* 2007; Yuan *et al.* 2011). In the case when secondary observables such as dispersion data are considered, the contribution from outside of the target region is calculated once and for all in the background global model, and subtracted from the observed dispersion data. The resulting residual is attributed to structure in the target region and the tomographic inversion proceeds within this target model volume. When inverting waveforms, and in particular when a more accurate theory than the PAVA is considered, such a simple procedure is not possible, and it is necessary to recompute the synthetic teleseismic waveforms, at each iteration, in a 3-D model which is fixed outside of the target region, and updated only within it. This leads to substantial computations, even in the case of an approximate theory such as NACT, and becomes prohibitively expensive, if the 3-D teleseismic wavefield is computed using SEM or another accurate numerical method, especially as one seeks to include increasingly shorter periods.

To overcome this problem, we take advantage of the 'Box Tomography' theory which allows us to combine teleseismic and regional waveforms in a consistent manner to image regional targets at arbitrary locations. This general framework accommodates arbitrary acquisition setups (i.e. with sources and receivers located outside the regional imaging box), is compatible with most popular numerical methods such as SEM or finite differences, and can produce accurate seismograms and sensitivity kernels through a complex and heterogeneous medium such as the Earth's mantle. In this study, we limit ourselves to the situation where both regional and teleseismic events are employed but where all the seismic stations lie within the regional computational domain. Furthermore, we neglect some higher order scattering effects (i.e. that correspond to waves that escape the local modeling domain, interact with the structure outside the modeling domain, and get scattered back inside the modeling domain), as proposed by Masson & Romanowicz (2017a). In this situation, Box Tomography is particularly efficient, as the computational effort to account for teleseismic data in regional inversions involves a limited number of global-scale simulations that are performed once and for all prior to the inversion. Masson & Romanowicz (2017a) showed, albeit in 2-D, that this approach can produce accurate regional tomographic images even though the elastic structure outside the imaged region is neither fully known prior to the inversion, nor updated during the inversion. Also, because we are working at relatively long periods, we assume that the effect of neglecting multipathing on teleseismic surface waves, that can be strong at 20 s and shorter, is not significant.

Here, we apply this methodology to the case of teleseismic three-component waveform data observed at stations within NA, combined with 'regional' waveform data for which both earthquakes and stations are located within the target region. This is the first time this methodology is applied to real three-component seismograms. Adding teleseismic data allows better azimuthal coverage of the target region than can be achieved using only the regional data set.

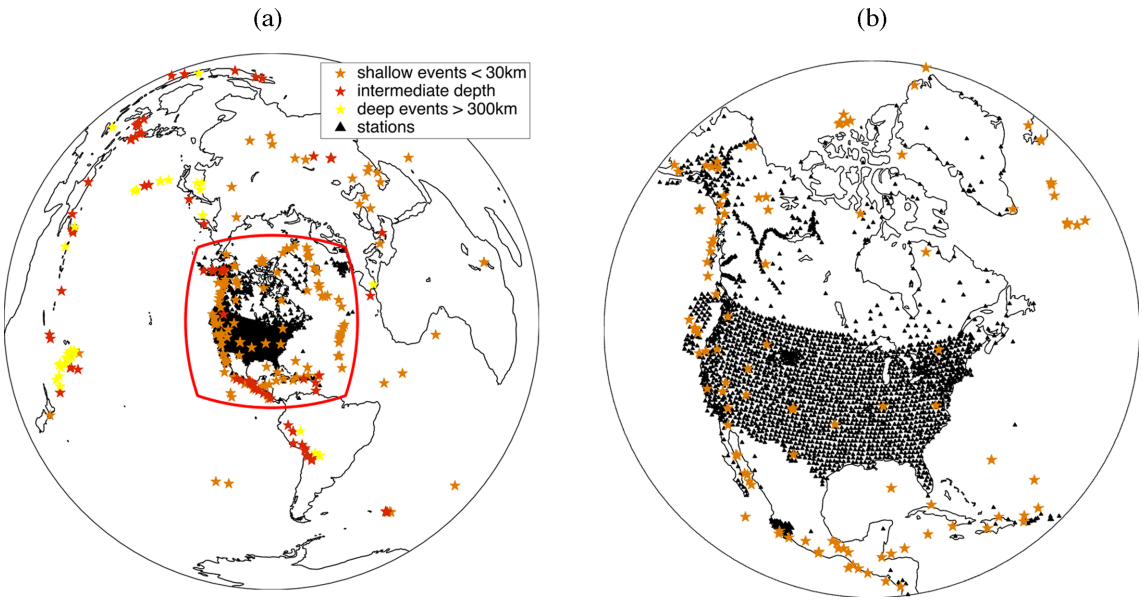

Figure 2. Source and station distribution considered in this study. Black triangles indicate seismic stations. Stars mark the location of 155 regional and 122 teleseismic events, with colour indicating depth. Thick red line indicates the boundaries used for the RegSEM forward simulation, which extends 89° × 89° horizontally down to the CMB. It contains the regional events and all the stations considered. (b) is a zoomed version of (a) to better show the distribution of stations.

2.1 Data set and model parametrization

.

The data set includes three-component acceleration waveforms from 2860 permanent and temporary broad-band stations located in the target region, a 89° × 89° area encompassing most of NA (Fig. 2). We considered two different data sets:

(i) a regional data set consisting of 155 events ($4.5 < M_w < 6.0$) for which sources are located within the target region.

(ii) a teleseismic data set, consisting of 122 events ($5.5 < M_w < 6.9$) for which sources are located outside of the target region.

The waveforms are filtered between 40 and 400 s, with corner frequencies at 53 and 250 s, and windowed into wavepackets, according to the procedure described in Li & Romanowicz (1996), allowing different weights to be applied according to relative amplitude of individual wavepackets, redundancy of paths and signal-to-noise ratio. When comparing our waveform data with the corresponding synthetics, two quantities are calculated for each selected wavepacket:

$$\mathrm{RMSR} = \frac{\sum_{i=1}^{N}(d_i - s_i)^2}{\sum_{i=1}^{N} d_i^2}$$

$$\mathrm{RMSS} = \frac{\sum_{i=1}^{N}(d_i - s_i)^2}{\sum_{i=1}^{N} s_i^2} \qquad (1)$$

where N is the number of data points in the wavepacket, d_i is the ith data point and s_i is the ith point in the synthetic trace. Wavepackets for which the value of either RMSR or RMSS is higher than 2 are rejected. A correlation coefficient is also calculated, and the wavepacket is rejected if it is negative. Several factors can contribute to the rejection of a wavepacket: poor signal-to-noise ratio, an incorrect instrument response, or the presence of glitches. Also, a wavepacket can be rejected if the current 3-D elastic model, through which its corresponding predicted waveform is calculated, provides too poor a fit to the waveform. The increase in the number of wavepackets retained as a function of model iteration provides an indication that the model is evolving in the right direction. A weighting scheme is then applied to the remaining wavepackets. The weight is a product of three terms $w_e \times w_n \times w_r$ where:

(i) w_e, is the inverse of the variance in the data, to equalize amplitudes across different wavepackets and different events.

(ii) $w_n = n^{-1/2}$, where n is the number of points in the wavepacket.

w_n aims to account for the redundancy among the data within a wavepacket.

(iii) $w_r = \frac{1}{P}$ quantifies the redundancy of the wavepackets with respect to wavepackets sampling similar ray paths. A radius of influence is defined around each source and station and P is the number of paths falling inside this influence zone (see fig. A1 in Li & Romanowicz 1996, for illustrations).

This weighting step is crucial, as it homogenizes the data coverage within the region and provides a way to construct the data covariance matrix C_D, which we approximate as a diagonal matrix. At each iteration of the inversion, only those wavepackets are considered that satisfy pre-defined goodness-of-fit criteria compared to synthetics calculated in the current 3-D model, in particular to avoid cycle skipping in our time-domain waveform inversion procedure. As the model improves, at each iteration, additional wavepackets are included as summarized in Table 1, while variance reduction increases, as shown in Fig. 3 and Table 1. The total number of wavepackets of fundamental and overtone Love and Rayleigh waves, is given in Table 1, and the path coverage within the NA continent including azimuthal coverage is given in Fig. 4.

The period range of our waveform data set is 40–400 s (broader than in Yuan *et al.* 2011, where the short cut-off period was 60 s for teleseismic data). In addition to long-period waveforms, we also consider a group velocity dispersion data set (Shapiro & Ritzwoller 2002) provided in the form of 1° × 1° maps between 25 and 100 s. The shorter periods (below 60 s) are used

Table 1. Total wavepacket counts (N_{wp}) and mean variance reductions (VR) in per cent for the model presented in this study and for the starting model, $SEMUCB\_wm1$, shown in brackets. The variance reduction is defined as $100 \times (1 - wr^2)$, where wr are the weighted residuals. Wavepackets are listed by component and data type (fundamental, higher and mixed modes). **F.** refers to fundamental modes, **M.** refers to mixed modes and **O.** refers to overtones. The increase in the number of wavepackets accepted is an indication of the improvement of the final model with respect to the starting model. It is particularly significant for *T*-component overtones and *Z*-component fundamental modes.

	L		*T*		*Z*	
	VR (per cent)	N_{wp}	VR (per cent)	N_{wp}	VR (per cent)	N_{wp}
F.	82.43 [76.04]	15887 [14454]	73.26 [61.08]	21651 [19844]	86.67 [76.5]	51397 [46395]
M.	82.40 [82.32]	13477 [13040]	72.20 [70.31]	20641 [19000]	80.52 [80.4]	29763 [28602]
O.	86.87 [79.54]	20738 [19726]	86.06 [62.54]	13031 [11679]	90.15 [79.1]	52787 [52371]

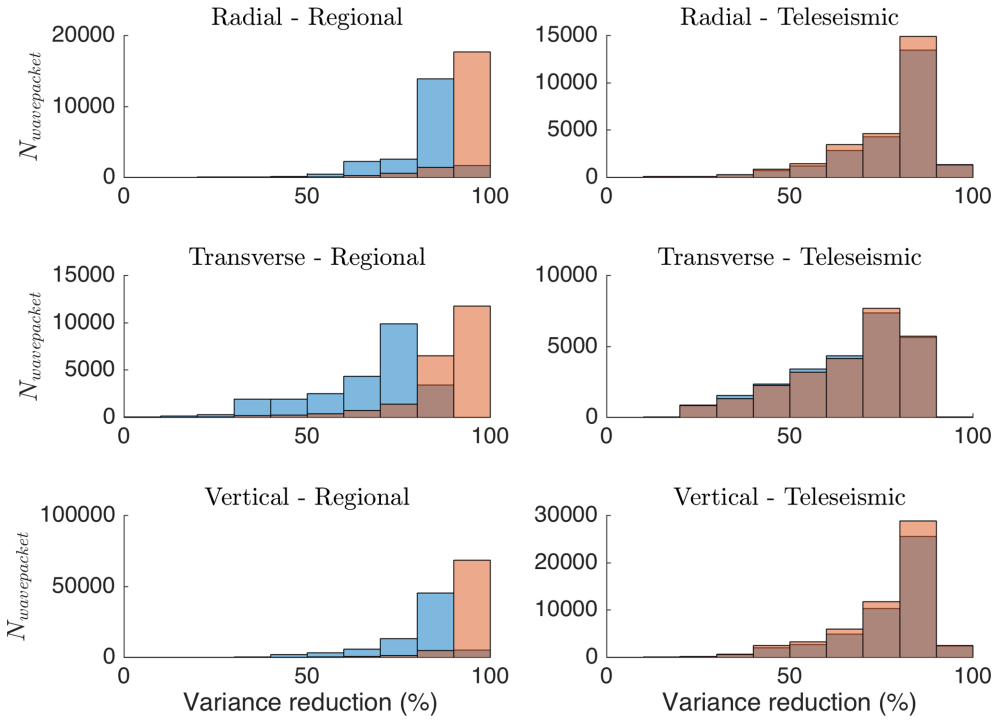

Figure 3. Histograms showing variance reduction between data and synthetics by component, separating the teleseismic and regional data sets. The *y*-axis gives the number of wavepackets and the *x*-axis is the variance reduction in per cent. The variance reduction is defined (as in Table 1 as $100 \times (1 - wr^2)$ where wr are the weighted residuals. The left-hand column shows variance reductions for regional events and the right-hand column for teleseismic (outside of RegSEM chunk) events. In each panel, blue/red bins show waveform variance reductions for the starting model and the current model, respectively. Dark red allows to see blue bins behind red bins. Variance reduction is much less significant for the teleseismic data, as the misfit is mostly due to the model outside of the region of study, which has been kept fixed.

to constrain our homogenized crustal structure at each iteration, as described below, while the entire period range of the dispersion data set is included during the inversion for mantle structure, providing additional constraints for structure in the shallow parts of the mantle, that are consistent with the treatment of the crust.

As in our previous work at the global scale (French *et al.* 2013; French & Romanowicz 2014), we do not rely on an existing layered crustal model. There are several drawbacks to considering such models. First, they are constructed using data for limited regions that are then extrapolated to other regions based on a tectonic regionalization, and may not always fit real waveform data very well. Second, the inclusion of thin low-velocity layers slows down the SEM computation significantly. Instead, we compute a radially anisotropic smooth crustal model on a $2^\circ \times 2^\circ$ grid through Monte Carlo Markov Chain (MCMC) simulation constrained by the group velocity dispersion data up to 60 s.

This process is called 'homogenization' (e.g. Backus 1962; Capdeville & Marigo 2008), and yields a crustal model equivalent to a real crust within the period range of the data used to constrain it (i.e. here for periods longer than 25 s). The thickness of the crust is *a priori* fixed to that of model Crust2.0 (Bassin *et al.* 2000) except in regions where Crust2.0 has a Moho at depths shallower than 30 km (i.e mostly in oceanic regions). In those regions, the crustal thickness is artificially fixed to 30 km. This results in a smooth, but realistic crust within most of the continent, and a crustal model that fits the data but is not directly interpretable in the border regions of our model space (i.e. in the oceans and borders of the continent where the real crust is thin). We have shown that this 'effective' model of the crust directly constrained by dispersion data does not bias the mantle structure imaged below 50 km (French & Romanowicz 2014). At each iteration of the inversion, we only invert for structure below 30 km depth, and subsequently recompute the effective crustal model for the next iteration by applying our

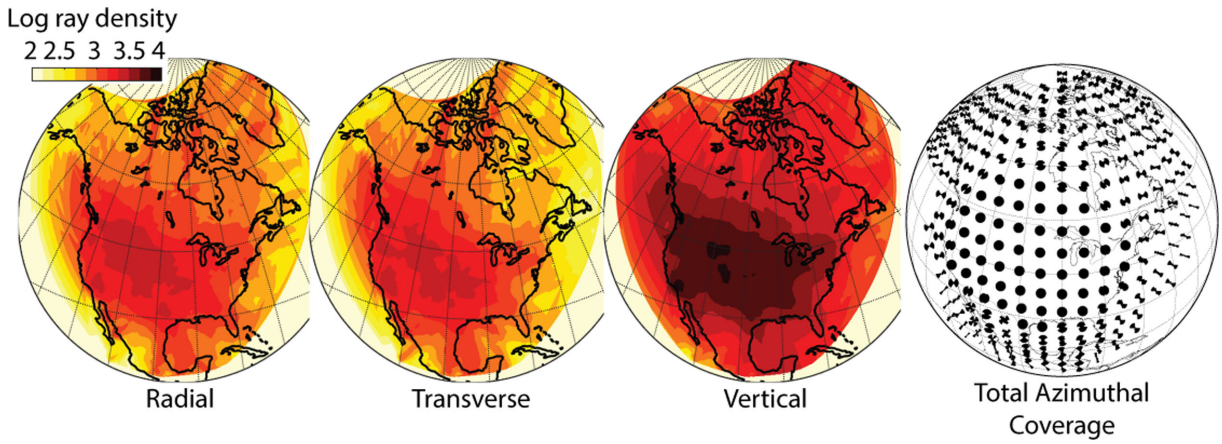

Figure 4. Density of paths and azimuthal coverage, projected at the earth's surface. The first three left-hand panels represent the density coverage for the different components. The region was divided into $2^\circ \times 2^\circ$ bins where the number of paths crossing the bin are counted and represented in $\log_{10}$ scale. The panel on the right shows the azimuthal coverage (rose diagram) for all components. Here, at each point on the map, the azimuthal range was divided into 5° circular bins. Each circle bin was cut into 15° slices. The number of paths crossing a slice is counted in a direction perpendicular to the circle edge. Each bin's radius is normalized to 1000 counts.

MCMC approach with mantle structure fixed to that of the current iteration model, in order to ensure the compatibility of the crust and uppermost mantle structure, which both affect the dispersion data used in the construction of the effective crustal model. The procedure is described and discussed in detail in French & Romanowicz (2014)

Assuming that our azimuthal coverage is everywhere sufficient within our target region, we here solve only for radially anisotropic structure. If A, C, F, L and N are the five elastic parameters describing a radially anisotropic (or VTI) medium (Love 1911), then:

$$N = \rho V_{SH}^2, \quad L = \rho V_{SV}^2, \quad A = \rho V_{PH}^2, \quad C = \rho V_{PV}^2 \tag{2}$$

where V_{SH}/V_{PH} and V_{SV}/V_{PV} are, respectively, the shear/compressional velocities of waves polarized horizontally and vertically and F represents intermediate orientations. The elastic medium can equivalently be described by the five parameters $(V_S, V_P, \xi, \phi, \eta)$ where:

$$V_S = \sqrt{\frac{2V_{SV}^2 + V_{SH}^2}{3}}, \quad V_P = \sqrt{\frac{V_{PV}^2 + 4V_{PH}^2}{5}} \tag{3}$$

$$\xi = \frac{N}{L}, \quad \varphi = \frac{C}{A}, \quad \eta = \frac{F}{A - 2L} \tag{4}$$

We only consider two of the five independent anisotropic parameters, those to which long-period surface waveforms are the most sensitive: isotropic velocity V_S and the anisotropic parameter $\xi = (\frac{V_{SH}}{V_{SV}})^2$. The other three parameters and density are constrained through empirical scaling relationships, following Montagner & Anderson (1989), and are based on laboratory measurements for upper-mantle rocks.

The scaling parameters considered are (Montagner & Anderson 1989)

$$\frac{\delta(V_P)}{\delta(V_S)} = 0.5, \quad \frac{\delta(\rho)}{\delta(V_S)} = 0.33, \quad \frac{\delta(\eta)}{\delta(\xi)} = -2.5, \quad \frac{\delta(\varphi)}{\delta(\xi)} = -1.5, \tag{5}$$

We chose a parametrization in terms of V_S and ξ, rather than V_{SH} and V_{SV}, as this allows us to consider different spatial resolution and apply higher damping in the inversion to the less well-resolved anisotropic parameter ξ, rather than having to reconstruct this parameter from differences in perturbations in the two quantities V_{SH} and V_{SV}, which would limit the resolution allowed in isotropic velocity.

In previous studies (Marone & Romanowicz 2007; Yuan & Romanowicz 2010), after inverting for radial anisotropic structure (Marone *et al.* 2007; Yuan *et al.* 2011), we also inverted for azimuthal anisotropy, by adding constraints from SKS splitting measurements. This step will be the topic of a separate publication.

The target model space is geographically defined as shown in Fig. 2. As in our previous tomographic studies, the global model space is parametrized in terms of cubic splines $v_q(r)$ vertically (Mégnin & Romanowicz 2000). In the target region, we invert for structure only in the top 700 km of the mantle, described by 16 cubic splines, for which the nodes are located at the following radii: 5690, 5810, 5910, 5900, 6061, 6101, 6131, 6161, 6191, 6221, 6241, 6261, 6281, 6301, 6321 and 6341 km. Laterally, we parametrize our model in terms of spherical splines $\beta_p(\theta, \phi)$ (Wang & Dahlen 1995).

The spherical spline parametrization has the advantage of allowing for variable grid parametrization, which can be adjusted according to data coverage (e.g. Marone *et al.* 2007). In our case, outside of the target region, we adopt a 'level 6' spherical grid for V_S (2° knot spacing) and a 'level 4' spherical grid for ξ (8° knot spacing), consistent with the parametrization of $SEMUCB\_WM1$ (French & Romanowicz 2014). Inside the well-sampled target region, we define a level 7 spherical grid (1° knot spacing) for V_S and level 6 (2° knot spacing) for ξ.

The combination of vertical and spherical splines constitutes a local basis for the description of smooth functions within the model volume. Thus, the value of a model parameter $m(r, \theta, \phi)$ can be computed at any point in space given the set of spline coefficients m_{pq}:

$$m(\theta, \varphi, r) = \sum_p \sum_q m_{pq} \beta_p(\theta, \varphi) v_q(r) \tag{6}$$

2.2 Forward modeling

During the inversion iterations, all the synthetic seismograms are computed using the SEM. The synthetic seismograms associated with regional data (i.e. where both the seismic stations and the earthquake are located within the regional modeling domain) do not

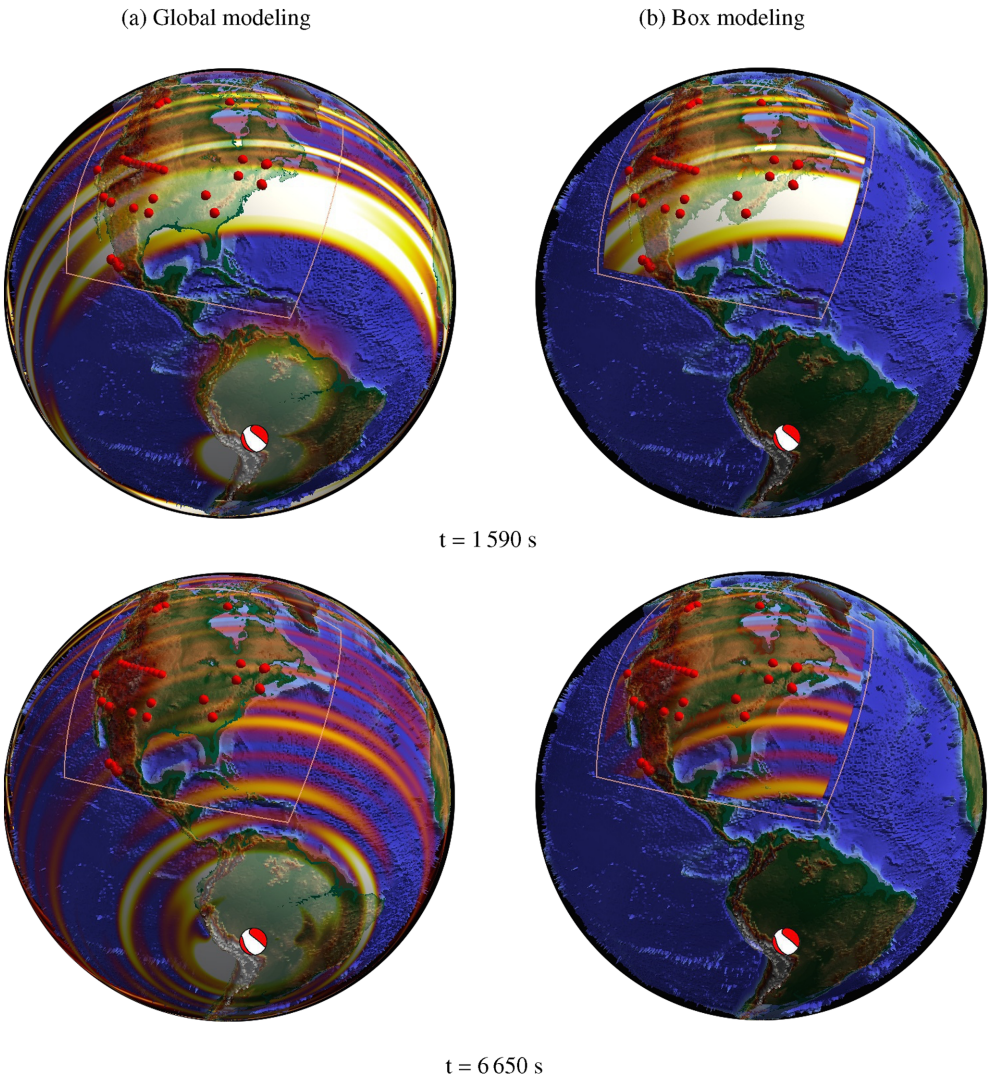

Figure 5. Comparison between a global scale and a regional (Box) scale numerical wave propagation simulation following an earthquake in south America. The 'beach ball' symbol shows the earthquake epicentre. In (a), the wavefield is modeled globally using the spectral element code SPECFEM_globe (Komatitsch 2002). In (b), the wavefield is modeled regionally using the spectral element code RegSEM (Cupillard *et al.* 2012), the regional wavefield is regenerated thanks to virtual sources placed at the border of the modeling domain (see: Masson *et al.* 2014). The top and bottom snapshots show the norm of the displacement vector at times 2590 and 6650 s, respectively. During subsequent iterations of the inversion, synthetics are computed using the regional solver RegSEM as shown in (b). Note that the dynamic range of the colour scale slightly differ between (a) and (b), the global and the local wavefields are however virtually identical as shown in Fig. 6.

require any specific treatment and are computed as in our previous studies using RegSEM (e.g. Yuan *et al.* 2014). The regional SEM code RegSEM (Cupillard *et al.* 2012) that takes into account effects of oceans, topography/bathymetry, ellipticity and anelasticity, and where the limits of the computational domain (orange box around NA in Fig. 5 and red box in Fig. 2) are dealt with using absorbing boundaries ('perfectly matched layers').

The synthetic seismograms associated with teleseismic data (i.e. where the earthquake and the seismic stations are located outside and inside the regional modeling domain, respectively) are obtained using a two-step procedure as proposed by Masson & Romanowicz (2017a) as follows.

Prior to the inversion, the wavefields generated by teleseismic earthquakes are computed globally within our starting model ($SEMUCB\_WM1$, which includes 3-D mantle structure and an associated homogenized crust, French & Romanowicz 2014) and a version of $SPECFEM3D\_globe$ (Komatitsch 2002) adapted to our model representation. During these global simulations, the three-component displacement wavefield is recorded at a set of points with locations prescribed by the RegSEM code on the borders of the target region. Within the regional solver RegSEM, these points are the collocation or Gauss–Lobatto–Legendre points belonging to the one-element-thick surface surrounding the regional modeling domain (see Masson *et al.* 2014). Note that our procedure does not require to store either the stress or the strain wavefield, which are computed naturally by the regional solver. This makes it easy to swap between different codes for modeling global seismic wave propagation (i.e. any code that outputs displacement seismograms can be used as such). Because the Courant criteria which ensure the stability and accuracy in SEM often lead to oversampling, we compress the recordings versus time using a least-squares B-spline transform (Unser *et al.* 1993a,b). We found this approach more efficient and practical than the more classic decimation/interpolation scheme. We typically achieve a data compression ratio of between

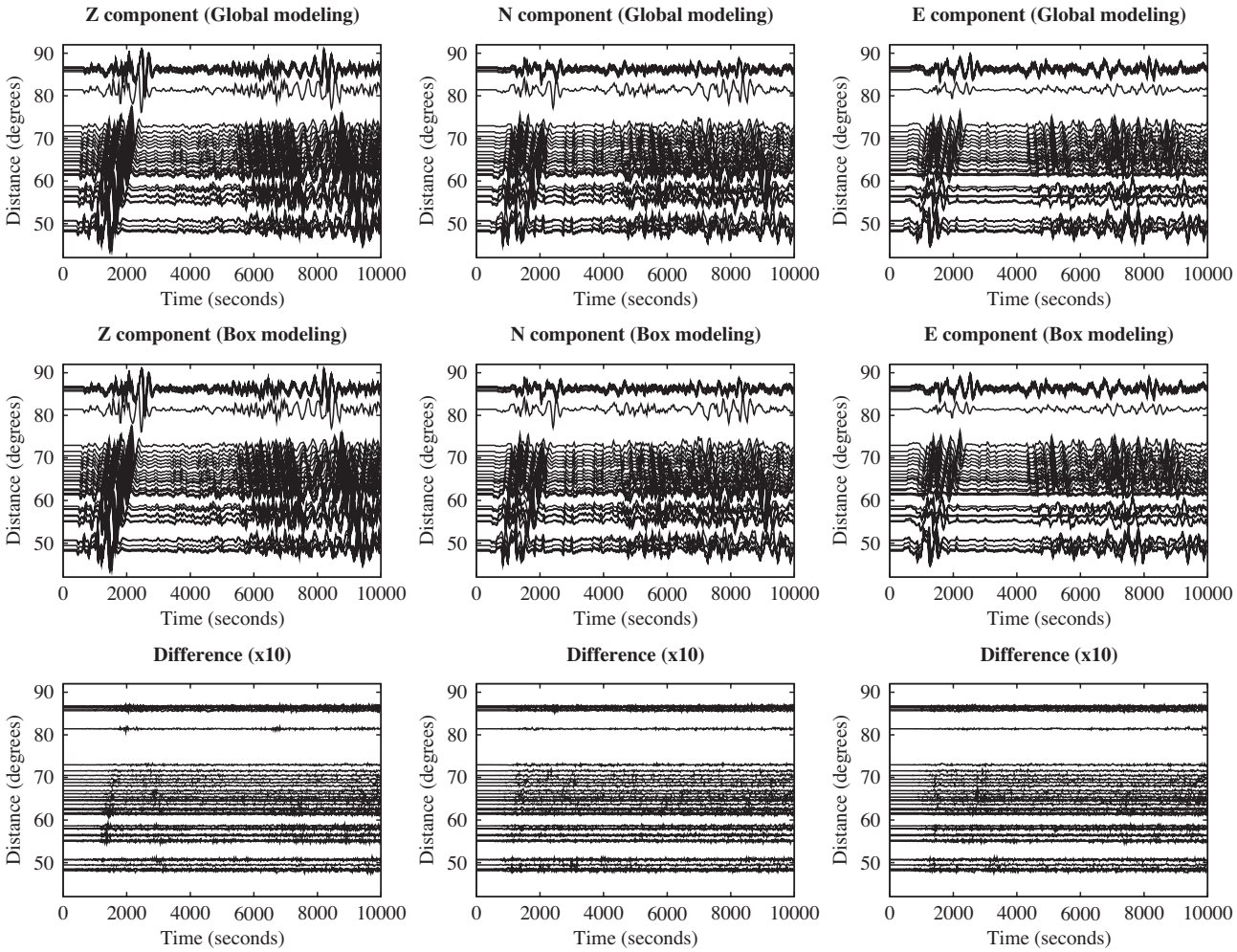

Figure 6. Comparison of synthetic seismograms filtered in the period range of 40–400 s. The top panels show three-component (ZNE) seismograms computed globally using the spectral element code SPECFEM3D_globe (Komatitsch 2002). The panels in the middle show the same using the spectral element code RegSEM (Cupillard *et al.* 2012). Those seismograms have been recorded during the simulations presented in Fig. 5. The three bottom panels show the difference between the global and the local seismograms; the difference have been multiplied by 10 and is plotted using the same scaling as the seismogram on the middle and top panels.

10 and 100 with no significant loss of accuracy. The global synthetic wavefield is thus transformed to virtual sources that regenerate the global wavefield regionally. This operation is done on the fly by the RegSEM code.

Overall, the additional computational effort to account for teleseismic events consists of one global simulation per teleseismic event. These simulations are performed once and for all before the inversion starts. A comparison between regular *SPECFEM3D_globe* and 'RegSEM + virtual sources' synthetics is shown in Figs 5 and 6. In particular, Fig. 6 shows that the difference between the global and regional computations is negligible in the period range considered. The total process time (CPU time) dedicated to the forward modeling in our Box Tomography inversion is given by:

$$\text{CPU\_box} = N_{\text{ext}} \times \text{CPU\_global} + N_{\text{iter}} \times N_{\text{tot}} \times \text{CPU\_regional} \quad (7)$$

where, N_{tot} is the total number of events, N_{ext} is the number of teleseismic events, N_{iter} is the number of iterations, CPU_global is the process time for one global simulation and CPU_regional is the process time for one regional simulation. As a comparison, the total process time of a full inversion performed the same way but where the forward modeling is done using global simulations only is:

$$\text{CPU\_full} = N_{\text{iter}} \times N_{\text{tot}} \times \text{CPU\_global}. \quad (8)$$

When the number of iterations is large enough, the ratio α between the process time of a full inversion and the process time of box inversion is simply the ratio between the process time of a global simulation and the process time of a regional simulation, we have:

$$\alpha = \frac{\text{CPU\_full}}{\text{CPU\_box}} \approx \frac{\text{CPU\_global}}{\text{CPU\_regional}}. \quad (9)$$

When modeling wave propagation using SEM, the process time to achieve one simulation is proportional to the volume V in which wave propagation is modeled. In general, the volume of the regional or box inversion V_{box} is much smaller than the volume of the Earth V_{Earth}. Thus, $\alpha \approx V_{\text{Earh}}/V_{\text{box}} \gg 1$ which shows that Box Tomography can be much more efficient than full inversions to image regional structures.

2.3 Inverse modeling

In continuity of our previous work (e.g Marone *et al.* 2007; Yuan *et al.* 2011, 2014), we use a hybrid iterative inversion scheme where, at each iteration, the forward wavefield is computed accurately using SEM, but the inverse step is solved using the formalism of Tarantola & Valette (1982) with sensitivity kernels calculated approximately using normal-mode perturbation theory. This allows us to apply a fast converging Gauss–Newton quadratic optimization scheme. As shown in appendix of Lekić & Romanowicz (2011), it is more important to use an accurate forward modeling scheme to compute the misfit function, provided of course that the theory captures the effects of the background long-wavelength structure accurately.

Inaccuracies in the theoretical treatment of kernels then result in smoothing errors that hopefully can be compensated for in subsequent iterations. Currently, popular 'adjoint tomographic' approaches (Zhu *et al.* 2015; Bozdağ *et al.* 2016) compute the misfit gradient numerically. However, since the numerical computation and storage of the full Hessian at each iteration is still prohibitive computationally (Chen *et al.* 2007b; Zhu *et al.* 2015, 2017), the misfit gradient is pre-conditioned by an approximate Hessian (usually diagonal). This step helps improve the convergence rate when solving the inverse problem (Pratt *et al.* 1998) which, generally in adjoint tomography, relies on a slow convergence linear optimization scheme (e.g. the conjugate gradient method, Zhu *et al.* 2015). In terms of regularization, the pre-conditioned gradient is smoothed using for example a 3-D Gaussian function (Tape *et al.* 2010). Such smoothing is gradually relaxed throughout iterations in regions with good coverage.

Our misfit function is defined in the time domain from the point by point differences between observed and synthetic waveforms as follows:

$$2\Phi(m_k) = [d - g(m_k)]^T C_d^{-1}[d - g(m_k)] + [m_p - m_k]^T C_m^{-1}[m_p - m_k] \quad (10)$$

where m_k represents the model estimate at the kth iteration, d is the data vector (waveform discretized in time or group velocity as a function of period) and $g(m_k)$ is the corresponding discretized wavefield computed using SEM, or the predicted group velocity dispersion. The model prior is m_p (i.e. the 3-D starting model) and C_m and C_d represent *a priori* model and data covariance matrices, respectively. Minimizing Φ in the sense of the L2 norm leads to the equation for the $(k+1)th$ model update:

$$m_{k+1} = m_k + (C_m G_k^T C_d^{-1} G_k + I)^{-1}(C_m G_k^T C^{-1}[d - g(m_k)] + m_p - m_k) \quad (11)$$

where G is the matrix of Frechet derivatives of $g(m)$ calculated at m_k. We compute G using NACT or PAVA, depending on the distance range of the corresponding source-station path. NACT and PAVA are both asymptotic (high frequency) approximations to first-order normal-mode perturbation theory. The PAVA includes along branch mode coupling only and is equivalent to the standard surface wave approximation (e.g Mochizuki 1986; Romanowicz 1987), as used for example in Woodhouse & Dziewonski (1984), in which a frequency and distance shifts are introduced for each mode to account for the effects of heterogeneous structure. The corresponding kernels are '1-D', that is, they only depend on the average structure between the source and the receiver. This approximation is most accurate for single-mode seismograms, such as fundamental-mode surface waves (e.g Romanowicz *et al.* 2008). NACT includes across branch coupling, in addition to PAVA, which brings out 2-D sensitivity of waveforms in the vertical plane containing the source and the receiver. However, NACT breaks down when the distance between the source and station is short, so we compute kernels using NACT for epicentral distances larger than 15°, and PAVA for shorter distances. Neither PAVA nor NACT consider off-great circle plane sensitivity [i.e. focusing effects, e.g. (Zhou *et al.* 2005)]. These effects become important for accurate amplitude fitting. With our choice of misfit function, we are first and foremost fitting the phase, and for that, the 2-D effects in the vertical plane are dominant, and important especially for overtones, as illustrated in Mégnin & Romanowicz (1999) and Romanowicz *et al.* (2008).

We first performed three iterations of simultaneous inversion for V_S and ξ, and found that the progression of the fit of transverse component was slower than for the vertical and radial components. We then decided to perform an additional iteration, inverting only for V_{SH} (while keeping V_{SV} structure fixed), followed by a final iteration inverting again for V_S and ξ simultaneously. The improvements in waveform fits between the fourth iteration and the last one are also captured in Table 2.

3 RESULTS

Average profiles of V_S and ξ with depth for different parts of the continent are presented in Fig. 7. The lateral variations in V_S and ξ are presented in map view at different depths in Figs 8 and 9, respectively, and in depth cross-sections in Figs 10 and 11, for V_S and Fig. 12 for ξ. In Figs A1–A6, we present the results of checkerboard tests which illustrate the resolution obtained for the adopted spatial parametrization in V_S and ξ, respectively, for the last iteration of the model. Each checkerboard test model is constructed by assigning a non-zero coefficient to a single radial spline $\nu_q(r)$ (eq. 6), resulting in 3-D structure concentrated around the corresponding depth. At the desired depth, the lateral structure is obtained by assigning the same value but opposite signs to each spherical spline $\beta_p(\theta, \varphi)$ (eq. 6) with 2° node spacing for V_S and 4° for ξ. The resolution matrix used in these tests is computed using the kernels and damping scheme of the last iteration of inversion with real data. A test of cross-contamination between V_S and ξ is shown in Figs A7 and A8, showing little contamination of the V_S structure by ξ structure. On the other hand, contamination of the ξ structure by heterogeneity in V_S is larger at depths greater than 200 km, but limited to regions of poor coverage. It is widely acknowledged that the standard resolution analysis employed in many tomographic studies is strictly valid only for linear problems (Tarantola 2005). Though waveform inversion is strongly non-linear, the argument can be made that it behaves pseudo-linearly under certain conditions, such as near the optimal model (Tarantola 2005). Acknowledging the above restrictions, the linear analysis still provides the best currently available tool for probing certain characteristics of our model that result from data quality/coverage and the influence of *a priori* information. The lateral resolution is good over most of the continental area, except in the most northern part, but, as always in this type of inversion, we note a significant loss of amplitude, due to the damping applied and some depth smearing in ξ. Also, in the most northern part, amplitude loss is more important due to poorer resolution. Contamination of ξ to V_S is negligible over the continent except in the northern part where the model is less constrained (see Fig. A7). Contamination of V_S to ξ is more important, especially at depths greater than 200 km where sensitivity to ξ is decreased, but still remains small ($d\ln(\xi) < 0.5$ per cent, see Fig. A8) .

Table 2. Total wavepacket counts (N_{wp}) and variance reductions (VR) in per cent for the model NASEM (fifth iteration), the fourth iteration, and for the starting model, $SEMUCB\_wm1$. The variance reduction is defined as in Table 1, but shown here combining fundamental mode and overtones for each component. The small change in residuals from fourth iteration to the fifth (final) iteration indicates that convergence has been reached.

	L		T		Z	
	VR (per cent)	N_{wp}	VR (per cent)	N_{wp}	VR (per cent)	N_{wp}
NASEM, fifth iteration	84.60	50102	75.35	55323	86.85	133947
Fourth iteration	84.54	59888	75.01	55286	86.80	133793
$SEMUCB\_wm1$, starting model	79.89	47220	66.41	50523	78.79	127368

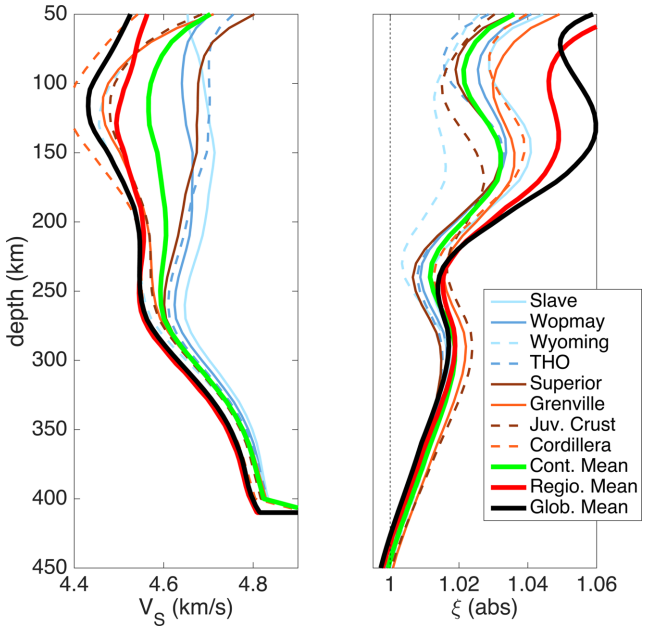

Figure 7. 1-D average depth profiles for isotropic shear velocity (left) and radial anisotropy (right), for the $SEMUCB\_wm1$ global model (thick black line), the continental average (thick green line) and for different regions in the model (see the legend and Fig. 1). The thick red line is the average within the region considered in the inversions, and is the reference profile for the maps shown in Figs 8 and 9. The continental average is always higher than the global average in V_S, while for ξ, they are similar at depths greater than 250 km.

We divide the model description and interpretation into several parts, first discussing the continent-wide structure and then, in more detail, that of the Eastern Superior craton.

3.1 Average velocity and anisotropy profiles with depth

As observed in many other studies, the average isotropic shear velocity V_S in the shallow upper mantle beneath the NA continent (green line in Fig. 7) is high compared to the global average of the model $SEMUCB\_WM1$, which includes contributions from large oceanic regions (black line in Fig. 7). The average over the region considered in the inversion is in-between (red line in Fig. 7). The difference between the continental and global average persists down to ∼270 km depth and then fades away, although V_S remains slightly higher in NA than in the global average down to 400 km depth, as observed in other models (e.g. Jordan & Paulson 2013). The Cordillera region shows the slowest velocities down to 200 km depth, lower than the global average, and more similar to those found in oceanic basins. Compared to the continental average, Precambrian provinces like Wyoming, Yavapai, Mazatzal, Great Rhyolites and Grenville exhibit slower velocities. Closer to the centre of the continent, regions such as the Slave, Trans Hudson Orogen (THO), Juvenile arcs and orogens (e.g. Wopmay) or Superior show higher velocities down to at least 250 km depth, with the Slave province uniquely characterized by a positive gradient with increasing depth above 100 km depth.

The continental average of ξ is larger than 1 (which means that $V_{SH} > V_{SV}$), but much lower than the global average, the latter influenced by the strong $\xi > 1$ oceanic signature. There are two maxima of ξ around 150 km depth and again around 300 km depth, as in the global average, although the first maximum is slightly deeper and laterally varying, and the second one is more pronounced. The anisotropic signature disappears at ∼400 km. The strongest variations relative to the continental average are observed at depths shallower than 200–220 km, (±4 to 8 per cent, compared to ±1 to 3 per cent at greater depths).

3.2 Upper-mantle structure within the craton

3.2.1 Lithospheric thickness

Our study confirms the presence of a thick lithospheric root in the craton with faster than average V_S down to 200–250 km depth, with highest velocities, ∼4.7–4.8 km s^{-1}, located at depths shallower than 150 km, mostly in the Superior, Rae and Slave provinces (e.g. Figs 8, 10 and 11). The lateral extent of the cratonic root, as defined by V_S larger than the continental average between 100 and 150 km depth, coincides roughly with the extent of Precambrian crust (Fig. 10, section A–A′). The western edge of the craton follows the RMF (black thick dashed line in Fig. 8), rather than the western continental rift boundary (grey dashed line in Fig. 8, which coincides with RMF in Canada, but diverges in the US). The eastern/southeastern edge of the craton coincides with the Llano-Grenville Front (red broken line in Fig. 8) rather than the 650 Ma eastern rift margin (eastern part of grey broken line in Fig. 8) as observed in Yuan *et al.* (2014) and Schaeffer & Lebedev (2014), except at shallow depths (60 km), where the extent of faster than average V_S reaches the 650 Ma eastern rift margin. To the north, the thick lithospheric roots extend to the Arctic ocean coastline and into Greenland.

While there are differences in the strength of the fast V_S anomaly among different parts of the craton at shallow depths, this becomes even clearer at larger depths, as will be discussed in the following sections.

3.2.2 Relation of lithospheric thickness to crustal age in the craton

As discussed in the previous section, the lateral boundaries of the craton, coinciding with the limits of Precambrian crust, are well observed down to 150 km depth. Below that, the lithospheric core retracts towards the centre of the continent comprising the following

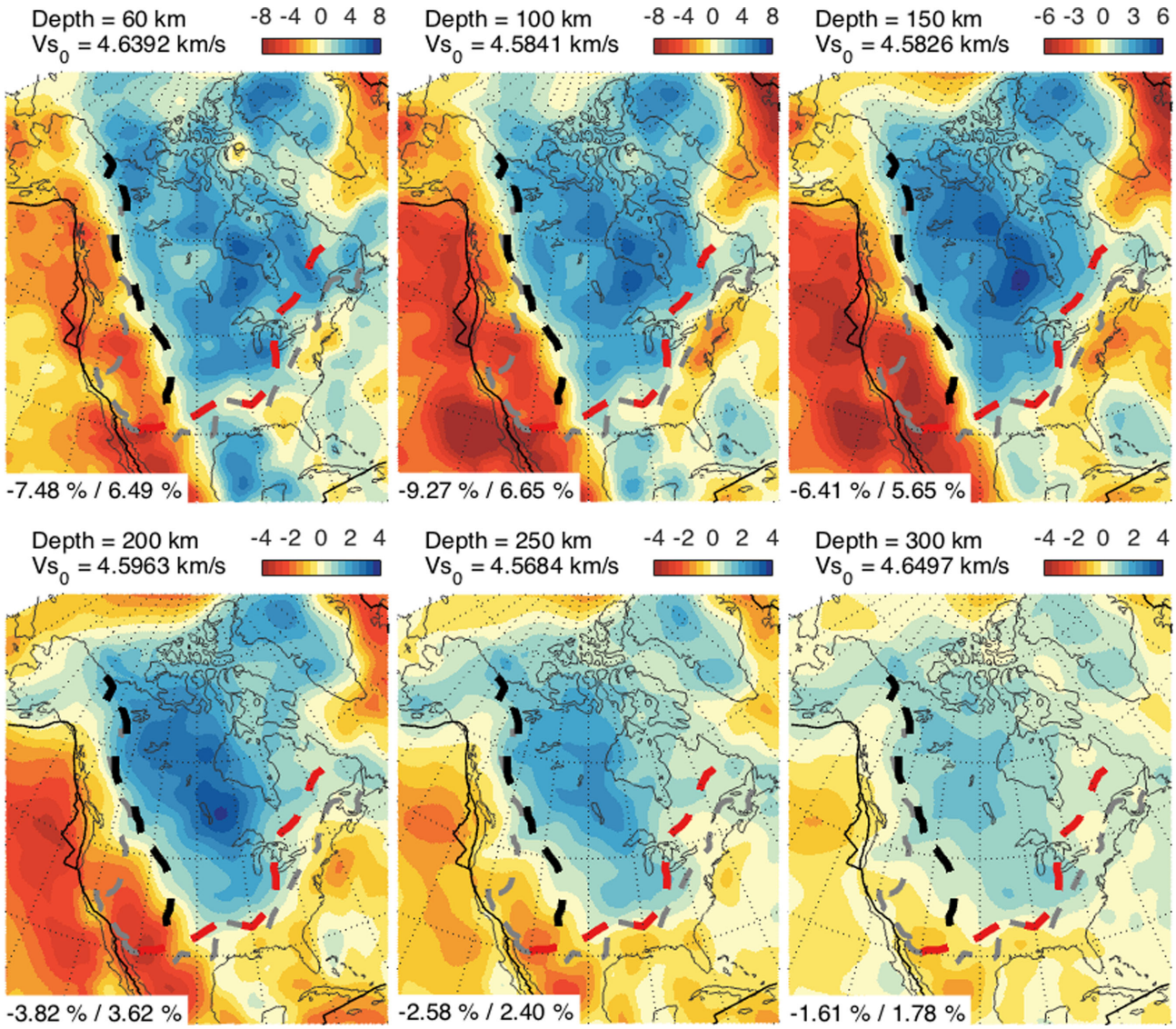

Figure 8. 3-D isotropic shear wave velocity structure of the continent shown in map view, as relative perturbations ($\frac{dV_S}{V_{S_0}}$) with respect to the regional mean (V_{S_0}, red curve in Fig. 7). Above each map, at the top left, depth and corresponding regional shear velocity average are indicated. Each colour bar, at the top right, indicates perturbations in per cent relative to the average value. Note that the colour scale is different at different depths. Minimum and maximum perturbations are given at the bottom left of each map. Black dashed line corresponds to the RMF, white thin lines show geological provinces as in Fig. 1, grey dashed lines indicate the western and eastern breakup lines and the red dashed line is the Grenville front.

provinces: north western Juvenile arcs and orogens (e.g. Wopmay and BHT), northeastern part of Juvenile provinces (e.g Yavapai, Mazatzal, Gr. Rhyolite), Slave, Hearne/Rae and western Superior cratons. In contrast, the eastern Superior Province and Wyoming craton disappear at depths greater than 150 km, as shown in Figs 8 and 10 (section B–B′). This is consistent with the global study of Bozdağ *et al.* (2016) but in contrast with recent regional (e.g. Nettles & Dziewoński 2008; Yuan *et al.* 2011, 2014; Schaeffer & Lebedev 2014) and global (e.g French & Romanowicz 2014, our starting model) models, where the core extends further to the northeastern part of the continent.

The mismatch between thickness of lithosphere and crustal age is in agreement with other studies in NA (Villemaire *et al.* 2012; Darbyshire *et al.* 2013; Porritt *et al.* 2015) and in several other cratons (e.g. Bruneton *et al.* (2004) for the Baltic shield, Fishwick *et al.* (2005) for the Australian shield), as well as the global cluster analysis of Lekić & Romanowicz (2011).

For example, the eastern Superior province of Archean crustal age shows a weak V_S anomaly below 150 km, in contrast to ∼+4 per cent V_S beneath the western Superior province (Figs 8 and 10, sections B-B′ and F–F′). The absence of a deep lithospheric root signature is also found in the Proterozoic Grenville province (Figs 8 and 10, sections C–C′, e.g. Villemaire *et al.* 2012; Bozdağ *et al.* 2016) and archaean Wyoming (Fig. 10, section D–D′, also observed in Yuan *et al.* 2014; Schaeffer & Lebedev 2014). On the other hand, the western part of the Proterozoic THO shows velocities at 150 km that are as high as found within the archaean Slave or western Superior lithosphere (see Fig. 11 with profile CC′).

Within the cratonic core, a maximum in V_S of ∼4.7–4.8 km s^{-1} is reached between 100 and 150 km (see Figs 10 and 11).

In Figs 10 and 11, we delineate the cratonic root by the contour of $V_S = 4.68$ km s^{-1} (dashed line). This is chosen somewhat arbitrarily in order to show how variable the depth of the cratonic root is across NA. It corresponds roughly to the middle of a zone of decreasing

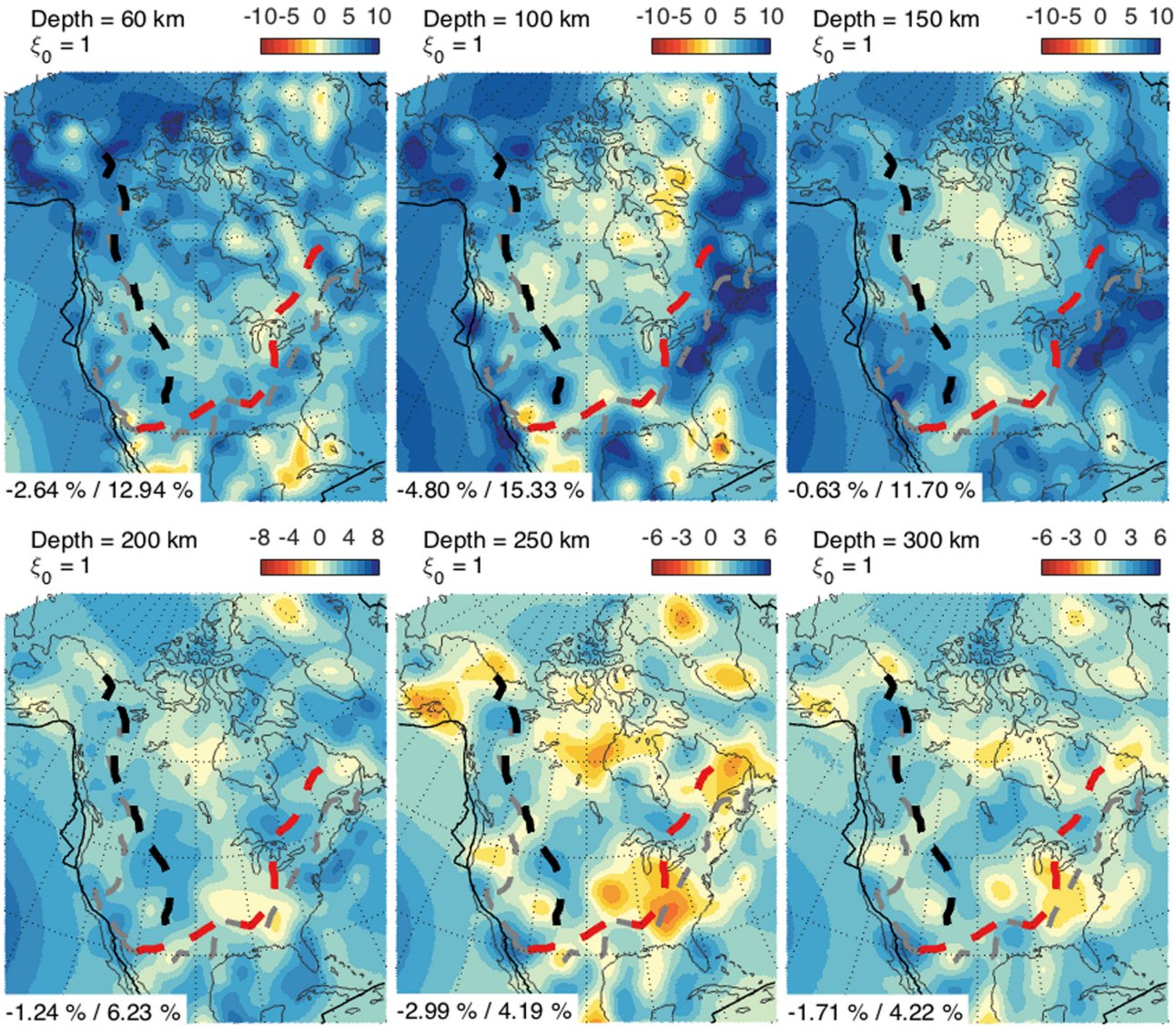

Figure 9. 3-D radial anisotropic structure of the continent shown in map view, as relative perturbations ($d\ln\xi = \frac{d\xi}{\xi_0}$) with respect to the isotropic case (e.g $\xi = 1$). Depth is indicated above each map, at the top left. Each colour bar, at the top right, indicates perturbations in per cent. The colour scale changes with depth. At the bottom left of each map, minimum and maximum perturbations are shown. Broken lines are as in Fig 8.

V_S which then reaches a minimum almost uniformly around 250–270 km depth, although there are lateral variations in the strength of the minimum (from less than 4.6 km s^{-1} to more than 4.65 km ;s^{-1}). Note that in tectonic areas of the western US, and in Baffin Bay in the north, the low-velocity minimum is located between 100 and 150 km depth, and is closer to the oceanic average.

Importantly, an LVZ is present everywhere under the craton, at the base of the lithosphere. We note that this feature was already present in our 3-D starting model. However, it is a very subtle feature under some of the cratonic regions, such as the Slave or the western Superior. Indeed the LVZ is very thin and velocity reductions are small (minimum 4.65 km s^{-1}) beneath regions where the lithosphere is thicker than 200 km. It is thicker and more pronounced under the eastern Superior province (minimum 4.58 km s^{-1}, e.g. Fig. 11 profile DD'). In the regions bordering the cratons, i.e. outside of RMF and Grenville Front, the low-velocity minimum is between 100 and 150 km and very pronounced ($V_S < 4.4$ km s^{-1}), and the transition out of it is more gradual, extending over ∼150 km depth, to reach velocities of 4.68 km s^{-1} at ∼300 km depth. On the other hand, the base of the LVZ, marked by a strong positive V_S gradient, is very smooth across provinces of all ages, as indicated by the 4.68 km s^{-1} line that oscillates between 300 and 325 km depth. Its minor undulations mirror the thickness of the lithosphere above, with slight upwarping beneath the deepest cratons, and downwarping under tectonic provinces, likely marking lateral variations in temperature extending to those depths.

Examining in more detail the various cratonic blocks, we see that, in the Slave province, the cratonic lithosphere is ∼200–250 km thick (Figs 8 and 10, cross-sections AA′ and DD′), but V_S is not as fast as within the southwestern/western Superior province. Down to 80 km, velocity anomalies are weak (1–2 per cent) compared to the entire Superior craton and Proterozoic belts (3–7 per cent), as is also the case in the northern Rae/Hearne cratons. This is highlighted by the $V_S = 4.68$ km s^{-1} contour line in Fig. 10, which appears around 80 km depth, indicating a positive gradient in velocity (as opposed to the negative gradient on the sides and at the base of the cratonic lithosphere everywhere beneath the craton). Still in the Slave province, the maximum of V_S around

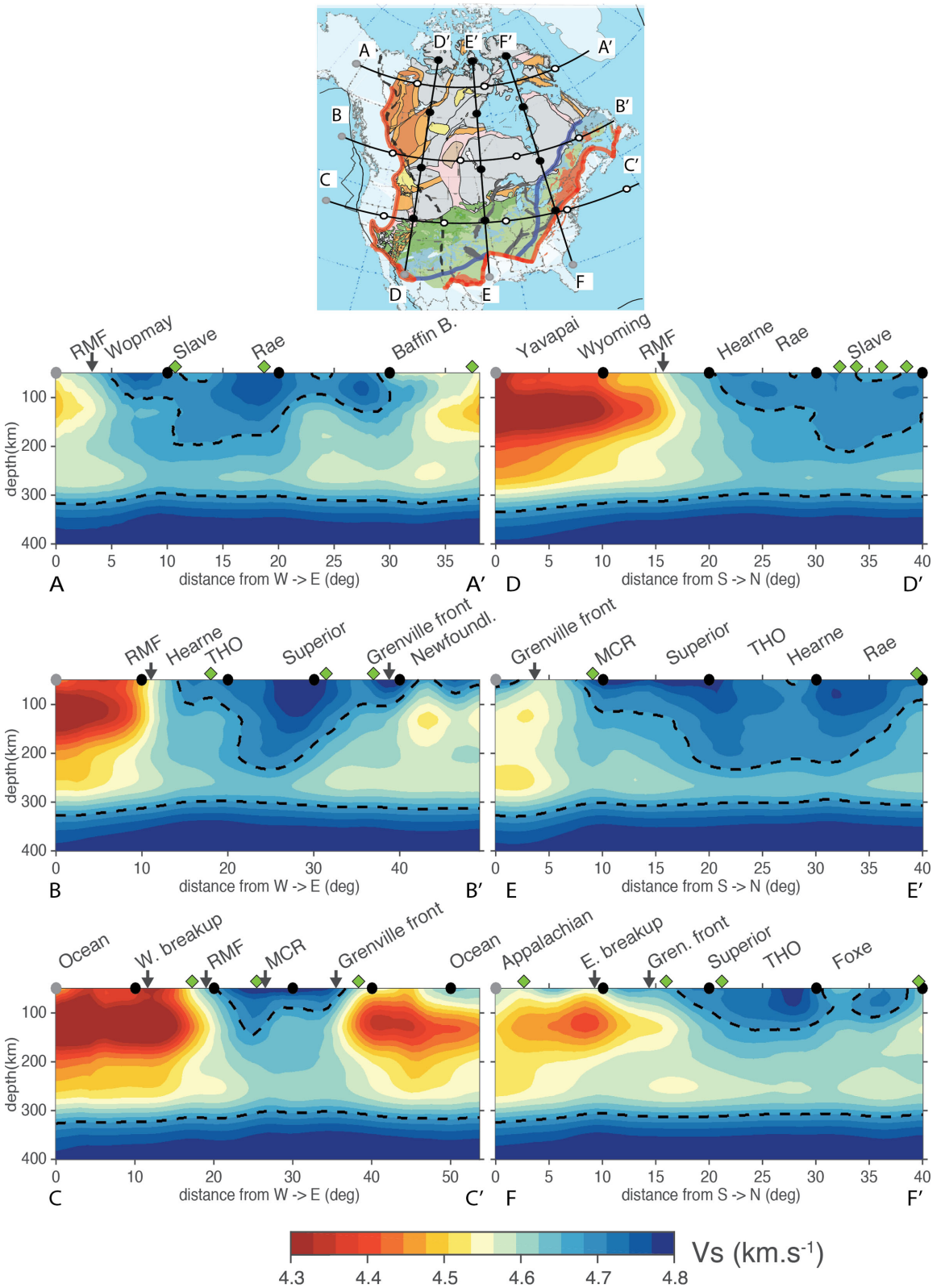

Figure 10. Depth cross-sections for V_S from 50 to 400 km deep. Absolute values of V_S are shown in km s^{-1} according to the colour bar at the bottom of the figure. The location of the different depth profiles shown is indicated on the geological map at the top. Small black/white circles on top of each cross-section correspond to the same black/white circles on the geological map. Green diamonds along each cross-section show the location of kimberlites from Faure *et al.* (2011) data set. Names indicate roughly which geological province is crossed (with a grey arrow to indicate a boundary, such as the western/eastern breakups, Grenville front, RMF). The black broken line marks the isocontour of V_S at 4.68 km s^{-1}.

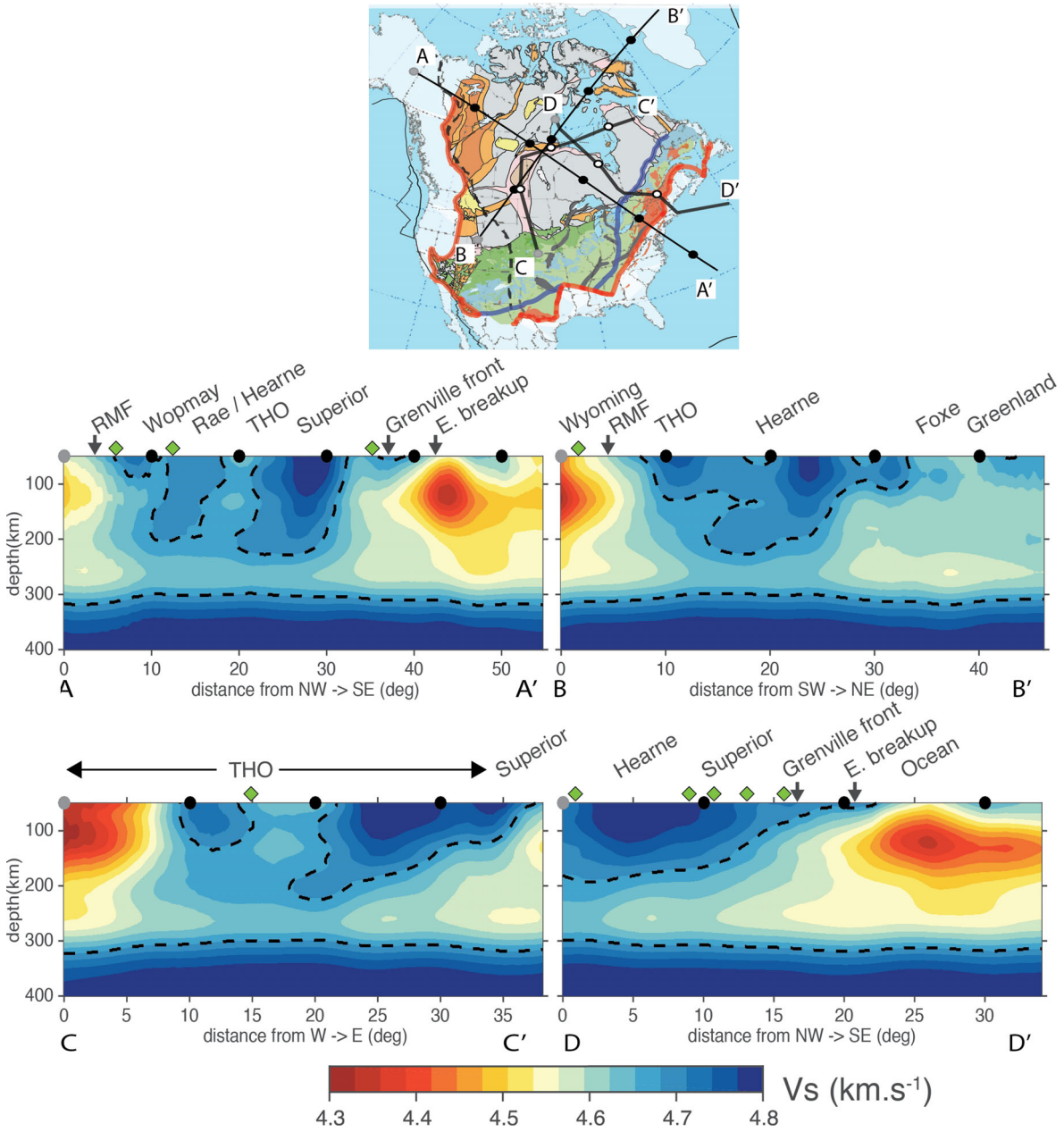

Figure 11. Same as Fig. 10 for four additional depth cross-sections in absolute V_S.

150 km depth does not exceed 4.7 km s^{-1}, a value slightly lower than in Chen *et al.* (2007a), but still within the error bar of 4.8 +/– 0.1 km s^{-1}. The relatively low velocities below 80 km depth may be the manifestation of the presence of an MLD, as documented for example in Snyder *et al.* (2014) and at station YKW3 in Calò *et al.* (2016).

In the northwestern part of the THO (beneath Hudson Bay), the lithosphere is more than 200 km thick (Fig. 11 profile CC′), while it is thinner in its eastern part (Fig. 11 profile BB′).

It is interesting to note that at depths shallower than 100 km, the western part of the THO, which welded the Superior and Rae/Hearne cratons (Hoffman 1988), and includes the Archaean Sask craton (Fig. 1) shows smaller (i.e. +4 per cent compared to +8 per cent) positive anomalies than in its Archaean surroundings (e.g. Fig. 11, profile AA′). However, at larger depths, velocities are higher and more similar to those of the surrounding Superior and Hearne regions. It has been suggested that the mantle may have been detached during the THO and replaced by both the Superior and the Hearne mantle lithosphere during final collision of the orogen (Németh *et al.* 2005; Faure *et al.* 2011).

However, a local study in Hudson Bay (Darbyshire *et al.* 2013) indicated the presence of a channel of slow velocities in THO compared to its surroundings down to 180 km depth. This is visible, but subtle, in our model (e.g. Fig. 11, profile AA′).

The southernmost extension of the craton (Yavapai/Mazatzal/Granit-Rhyolite orogens) represents zones of accretion of Proterozoic material and marks the amalgamation of the Archean cores (Hoffman 1988). It shows a weaker but deep root between the RMF and the Grenville front (cross-sections EE′ and CC′ in Fig. 10 and maps at 200 and 250 km depth in Fig. 8) with velocities reaching almost 4.7 km s^{-1} at ∼150 km depth. This deep feature is also found in the V_P model of Burdick *et al.* (2017). There is some indication

of incursion from rifting separating these younger crustal regions (Yavapai) at their northern edge from the Superior craton, manifested by lower velocities in the depth range 100–250 km (Fig. 10, profiles CC′ and EE′). Note that the high-velocity anomaly is located where cretaceous-aged kimberlites are present (also shown in Fig. 13). The lithosphere thickness derived from kimberlite samples is about 160 km, a depth that corresponds to the maximum in velocity in our model (Fig. 10, profiles CC′ and EE′). The cratonic root may have been affected by asthenospheric upwelling related to the Mid-Continent Rift System. (Griffin *et al.* 2004).

3.2.3 Radial anisotropy structure

As observed in Babuška *et al.* (1998), Gung *et al.* (2003), or Yuan *et al.* (2014), there are large lateral variations of ξ down to ∼300 km depth, with localized regions of reduced radial anisotropy, in contrast to the more uniform, $d\ln(\xi) > 0$ structure in this depth range in the oceans (see Fig. 9). The regions with $\xi < 1$ (i.e. $d\ln(\xi) < 0$) observed between 60 and 150 km depth include Mexico, the southern part of Florida, Baffin Island and Hudson Bay. Below 150 km $\xi < 1$ regions are observed in the southern part of the Superior craton, such as the Mid-Continental rift, the southern part of the Granite Rhyiolite Province comprising the failed rifting area, a large area comprising Hearne/Rae, Hudson Bay, eastern Slave, Greenland, Alaska and the southwestern part of the RMF (see Figs 12 and 9).

One of these regions, in the west, may be related to the Yellowstone hotspot, and another one to the mid-continent rifting episode, although similar explanations cannot be readily invoked in other parts of the region exhibiting $\xi < 1$. We note that such regions of $\xi < 1$ were not present in the starting model.

At 250 km depth, alternating zones of $d\ln\xi > 0$ and $d\ln\xi < 0$ are observed everywhere, with a band of stronger positive $d\ln\xi$ across the THO potentially reflecting higher deformations. Similar to Zhu *et al.* (2017), we observe $\xi < 1$ in the western US, however our model differs from the latter in the eastern US, where Zhu *et al.* (2017) find a strong $d\ln\xi > 0$ and we find $d\ln\xi < 0$ (profile CC′ in Fig. 12).

A minimum in ξ is observed everywhere around 250 km depth (Fig. 12), with some lateral variations in strength, in particular a more pronounced minimum beneath regions in the craton that exhibit thinner lithospheric roots, as defined by V_S: for example under Hearne (Fig. 12, profiles DD′ and BB′) and under the eastern Superior province (Fig. 12, profiles BB′ and FF′). The depth of this minimum ∼220–230 km seems shallower than the minimum in the V_S profiles, while the zone of increased anisotropy below that minimum peaks around 300 km depth, near the base of the LVZ in V_S.

Comparing panels BB' and EE' in Figs 10 and 12, we note, interestingly, that regions of marked $d\ln\xi < 0$ at depths greater than 200 km often correspond to regions of reduced anisotropy at shallower depths, which themselves correspond to borders of individual cratons (e.g. between Hearne and RMF in panels BB', or under MCR and Rae in panels EE'). This may indicate a connection between upwelling regions in the asthenosphere, and zones of weakness in the continental lithosphere. A similar observation was made in Yuan *et al.* (2011, see fig. 11 b in that paper).

To summarize, we find a lithospheric core beneath the Precambrian shield, as observed in many previous tomographic studies of NA, characterized by high V_S and $\xi > 1$, but the anisotropy is weaker than under ocean basins. We do not observe a direct link between the age of the crust and its underlying lithosphere everywhere. Most of the cratonic core is located beneath Precambrian crust, but the history of the amalgamation and destruction [e.g. convective removal, Lee *et al.* (2011)] of the craton has influenced the present structure of its lithosphere. This is the case for the Archean Wyoming craton, which does not share the characteristics of other Archaean blocks (e.g. Slave, Hearne, Rae and western Superior) and may have suffered from basal erosion (King 2005; Lee *et al.* 2011). Also, the lithospheric root in the eastern part of the Archean Superior craton shows a weaker signature. In the next section, we will focus on this craton, which may have endured basal deformation and weakening, induced by the passage of the Great Meteor hotspot track.

3.3 The Great Meteor hotspot track and lithospheric thickness under the eastern Superior province

As in Schaeffer & Lebedev (2014), Yuan *et al.* (2014), James *et al.* (2014), Silveira & Stutzmann (2002) and Mocquet & Romanowicz (1990), on the eastern border of NA, we observe slower V_S anomaly structure extending continuously between the Grenville front and the Atlantic coast down to 200 km depth (Fig. 8). Around latitude 40°N, this low-velocity anomaly continues further to the east, all the way to the mid-atlantic ridge, following the corresponding low-velocity 'finger' of French *et al.* (2013), and including the Great Meteor Track (GMT) and the Bermuda track. This is particularly clear at 150 km depth (Fig. 13). On the other hand, to the west, the slow velocity band continues across the Appalachian and Grenville orogens, and crosses the Superior Craton northward up to Hudson Bay, following the GMT (Heaman & Kjarsgaard 2000), at depths between ∼150 and 300 km (Fig. 11, profile DD′). This likely represents the thinning of the cratonic lithosphere associated with the passage of the continent over the Great Meteor hotspot during the opening of the Atlantic Ocean. Its location corresponds to the V-shaped dent into the craton observed by van der Lee & Nolet (1997), who, alternatively, suggested this feature may be due to the presence of water in the mantle. It is likely that in this eastern part of the Superior Craton, the deepest part of the lithosphere has been eroded due to the interaction of the hotspot, and following its passage. Xenoliths analysis from kimberlites at Kirkland Lake (southeastern part of the western Superior) exhibit some evidence of diamonds, suggesting that the base of the lithosphere was deeper than 170 km at the time of kimberlitic eruption (Jones *et al.* 2014) estimated at 160–152 Ma by Heaman & Kjarsgaard (2000) and Heaman *et al.* (2003). This is also supported by the study of Faure *et al.* (2011), which reports that 23 per cent of kimberlite samples from the southern Superior region contain diamonds, in contrast to 78 per cent in the northern Superior. Other studies propose similar scenarios: Davies (1994) described a process of thermomechanical erosion which involves heating–softening of the lithospheric base by a plume tail, followed by mechanical removal of the material due to convective transfer (Rondenay *et al.* 2000). Kaban *et al.* (2015) proposed that basal drag induced by plate motion, and facilitated by reheating of the lithosphere due to the passage of the Great Meteor hotspot, caused a southwestern shift of the Superior's cratonic root. A feature that was observed by the local inversion of gravity, topography, crustal structure and seismic data (e.g. Eaton & Frederiksen 2007; Kaban *et al.* 2015).

Our observations show that despite the Archaean age of the crust, the lithospheric root thickness of a craton can be altered by subsequent magmatism, as broadly observed within the eastern part of the Superior craton, similarly to what has been proposed for the China craton (e.g. Chen *et al.* 2009).

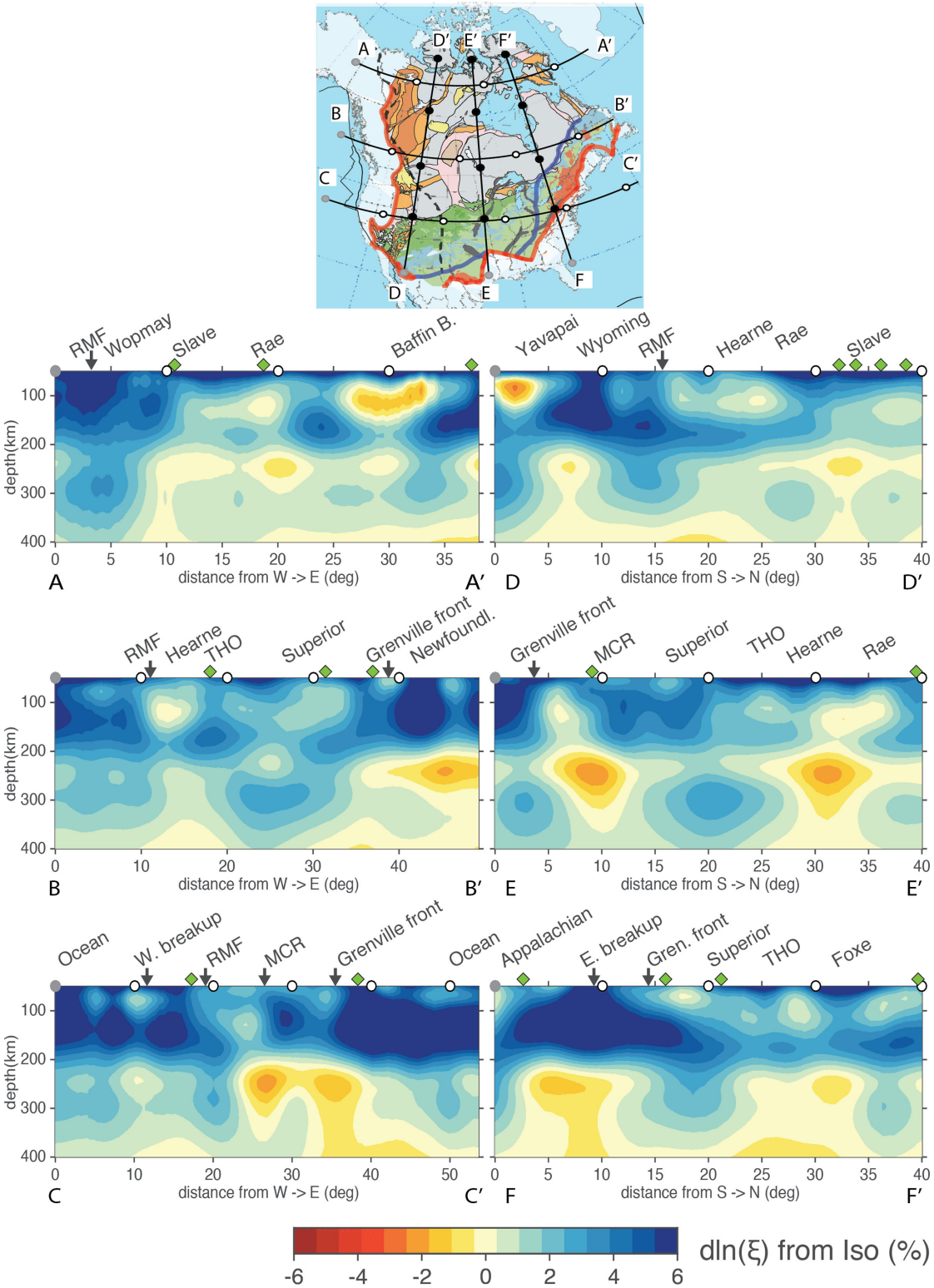

Figure 12. Same as Fig. 10 for ξ, shown here in per cent perturbation with respect to the isotropic case (i.e $\xi = 1$).

4 DISCUSSION AND CONCLUSION

We have developed a 3-D radially anisotropic model of the upper mantle under the NA continent, based on a combination of teleseismic and regional distance waveform data, consistently and seamlessly treated using numerical wavefield computations, as a first application of 'Box Tomography' in a case where stations are inside the target region, but sources can be both inside and outside the region. Our model shows well-defined craton borders coinciding in the west with the RMF, and in the east, with the Grenville Front and highlights lateral variations in lithospheric thickness within the NA craton that do not correspond to crustal ages everywhere. In particular, it confirms the difference in thickness of the eastern and

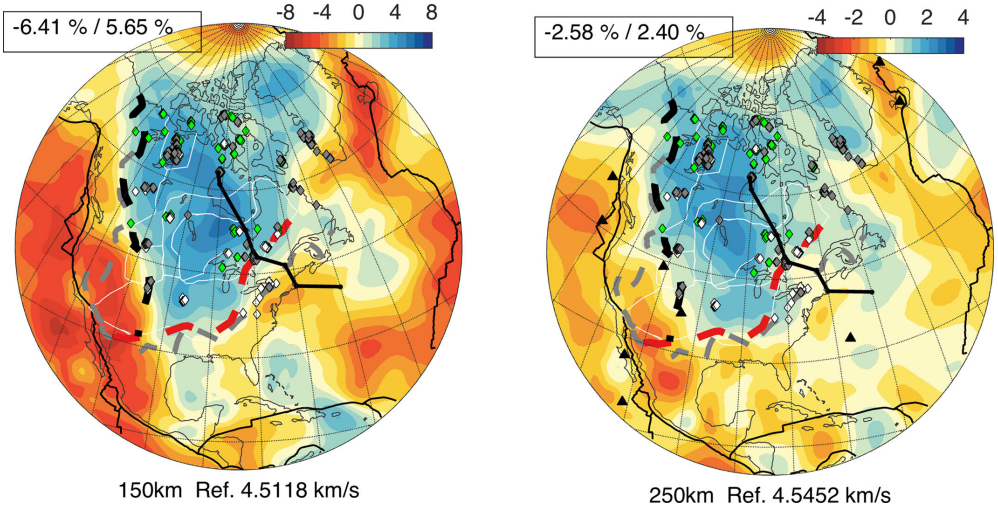

Figure 13. 3-D isotropic shear wave velocity map of the continent at 150 and 250 km depth, in terms of variations with respect to the regional mean, as $\frac{dV_S}{V_{S_0}}$. Black line corresponds to the Great Meteor Track from Heaman & Kjarsgaard (2000), black triangles show positions of hotspots and white, black, grey squares show kimberlites with/without/unknown presence of diamonds (Faure *et al.* 2011). Below the map, depth and the corresponding regional shear velocity mean is indicated. At the top left, minimum and maximum V_S perturbations are given.

western parts of the Superior province, and attributes the thinner lithosphere in the eastern part to the aftermath of interaction with the Great Meteor hotspot at the time of the last opening of the Atlantic Ocean. We also show the presence of a LVZ centred around 270 km depth everywhere in the craton, but with laterally varying topography and strength: it is less pronounced beneath regions of lithosphere thicker than 200 km, and wider and more pronounced in regions of thinner lithosphere. In contrast, the base of the LVZ shows smaller lateral variations in depth, with some upwarping under thick lithospheric roots, and downwarping in tectonic regions and regions of thinner lithosphere beneath the craton. The lateral variations in radial anisotropy structure also show some correspondence with the variation in the thickness of the lithosphere. In the depth range 60–150 km, two of the zones of weaker anisotropy within the craton may be associated with the Yellowstone and Great Meteor hotspots, respectively. In fact, regions of weaker anisotropy in the lithosphere at the edges of individual cratons are often associated with strong negative anomalies ($d\ln\xi < 0$) at depths 200 km and larger, suggesting the presence of asthenospheric upwellings in correspondence with zones of weakness in the continental lithosphere.

This study being foremost a feasibility study for 'Box Tomography', it is clear that resolution of the model could still be improved in the future. The remaining unexplained variance in the observed waveforms likely results from the fact that we did not update the source parameters of the events used in this study, nor did we consider azimuthal anisotropy, focusing/defocusing effects or lateral variations in attenuation. Another source of remaining misfit may come from the uncertainties in the model outside of the region which has not been updated in the iterative process presented here, although Masson & Romanowicz (2017b) have shown, at least in 2-D, that, with good coverage of the region, this is not a serious concern. Further improvement in the coverage of the region could be achieved by including stations outside of the "box" and decreasing the minimum period of the waveforms considered. This involves the numerical computation of Green"s functions between the stations and the box (e.g. Masson & Romanowicz 2017a), and will be implemented in the near future. Also, to increase resolution, the path-average approximation used in the computations of the inverse kernels needs to be replaced by more accurate theory. All these factors are likely to contribute to the observed waveforms and will be addressed in upcoming studies.

ACKNOWLEDGEMENTS

PC thanks Huaiyu Yuan and Scott French for introducing him to the waveform modeling codes used in this study. This research was supported by the European Research Council, under the European Community's FP7 Programme (advanced grant WAVETOMO). Numerical computations were partly performed on the S-CAPAD platform, IPGP, France as well as resources at the National Energy Research Scientific Computing Centre. BR also acknowledges support from NSF grant EAR-1460205.

REFERENCES

Abt, D.L., Fischer, K.M., French, S.W., Ford, H.A., Yuan, H. & Romanowicz, B., 2010. North american lithospheric discontinuity structure imaged by ps and sp receiver functions, *J. geophys. Res.,* **115**(B9), doi:10.1029/2009JB006914.

Babuška, V., Montagner, J.-P., Plomerová, J. & Girardin, N., 1998. Age-dependent large-scale fabric of the mantle lithosphere as derived from surface-wave velocity anisotropy, in *Geodynamics of Lithosphere and Earth's Mantle,* pp. 257–280, ed. Birkhuser, B., Springer.

Backus, G.E., 1962. Long-wave elastic anisotropy produced by horizontal layering, *J. geophys. Res.,* **67**(11), 4427–4440.

Bassin, C., Laske, G. & Masters, G., 2000. The current limits of resolution for surface wave tomography in north america, *EOS Trans Am. geophys. Un* **81,** F897.

Bedle, H. & van der Lee, S., 2009. S velocity variations beneath North America, *J. geophys. Res.,* **114**(B7), doi:10.1029/2008JB005949.

Bostock, M., 1998. Mantle stratigraphy and evolution of the slave province, *J. geophys. Res.,* **103**(B9), 21 183–21 200.

Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N. & Pugmire, D., 2016. Global adjoint tomography: first-generation model, *Geophys. J. Int.,* doi:10.1093/gji/ggw356.

Bruneton, M. *et al.*, 2004. Complex lithospheric structure under the central Baltic shield from surface wave tomography, *J. geophys. Res.,* **109**(B10), doi:10.1029/2003JB002947.

Burdick, S. *et al.*, 2014. Model update January 2013: upper mantle heterogeneity beneath North America from travel-time tomography with global and USArray transportable array data, *Seismol. Res. Lett.,* **85**(1), 77–81.

Burdick, S. *et al.*, 2017. Model update May 2016: upper-mantle heterogeneity beneath North America from travel-time tomography with global and usarray data, *Seismol. Res. Lett.,* **88**(2A), 319–325.

Calò, M., Bodin, T. & Romanowicz, B., 2016. Layered structure in the upper mantle across North America from joint inversion of long and short period seismic data, *Earth planet. Sci. Lett.,* **449,** 164–175.

Capdeville, Y. & Marigo, J.-J., 2008. Shallow layer correction for spectral element like methods, *Geophys. J. Int.,* **172**(3), 1135–1150.

Chen, C.-W., Rondenay, S., Weeraratne, D.S. & Snyder, D.B., 2007a. New constraints on the upper mantle structure of the Slave craton from Rayleigh wave inversion, *Geophys. Res. Lett.,* **34**(10), doi:10.1029/2007GL029535.

Chen, L., Cheng, C. & Wei, Z., 2009. Seismic evidence for significant lateral variations in lithospheric thickness beneath the central and western North China Craton, *Earth planet. Sci. Lett.,* **286**(1), 171–183.

Chen, P., Zhao, L. & Jordan, T.H., 2007b. Full 3d tomography for the crustal structure of the los angeles region, *Bull. seism. Soc. Am.,* **97**(4), 1094–1120.

Cleary, J. & Hales, A.L., 1966. An analysis of the travel times of P waves to north american stations, in the distance range 32 to 100, *Bull. seism. Soc. Am.,* **56**(2), 467–489.

Cupillard, P., Delavaud, E., Burgos, G., Festa, G., Vilotte, J.-P., Capdeville, Y. & Montagner, J.-P., 2012. RegSEM: a versatile code based on the spectral element method to compute seismic wave propagation at the regional scale, *Geophys. J. Int.,* **188**(3), 1203–1220.

Darbyshire, F.A., Eaton, D.W. & Bastow, I.D., 2013. Seismic imaging of the lithosphere beneath hudson bay: episodic growth of the laurentian mantle keel, *Earth planet. Sci. Lett.,* **373,** 179–193.

Davies, G.F., 1994. Thermomechanical erosion of the lithosphere by mantle plumes, *J. geophys. Res.,* **99**(B8), 15709–15722.

Debayle, E. & Ricard, Y., 2012. A global shear velocity model of the upper mantle from fundamental and higher Rayleigh mode measurements, *J. geophys. Res.,* **117**(B10), doi:10.1029/2012JB009288.

Deschamps, F., Lebedev, S., Meier, T. & Trampert, J., 2008. Stratified seismic anisotropy reveals past and present deformation beneath the east-central United States, *Earth planet. Sci. Lett.,* **274**(3–4), 489–498.

Eaton, D.W. & Frederiksen, A., 2007. Seismic evidence for convection-driven motion of the North American plate, *Nature,* **446**(7134), 428–431.

Faure, S., Godey, S., Fallara, F. & Trépanier, S., 2011. Seismic architecture of the Archean North American mantle and its relationship to diamondiferous kimberlite fields, *Econ. Geol.,* **106**(2), 223–240.

Fischer, K.M., Ford, H.A., Abt, D.L. & Rychert, C.A., 2010. The lithosphere-asthenosphere boundary, *Annu. Rev. Earth Planet. Sci.,* **38**(1), 551–575.

Fishwick, S., Kennett, B. & Reading, A., 2005. Contrasts in lithospheric structure within the Australian craton-insights from surface wave tomography, *Earth planet. Sci. Lett.,* **231**(3), 163–176.

Ford, H.A., Long, M.D. & Wirth, E.A., 2016. Midlithospheric discontinuities and complex anisotropic layering in the mantle lithosphere beneath the Wyoming and Superior Provinces, *J. geophys. Res.,* **121**(9), 6675–6697.

French, S., Lekic, V. & Romanowicz, B., 2013. Waveform tomography reveals channeled flow at the base of the oceanic asthenosphere, *Science,* **342**(6155), 227–230.

French, S.W. & Romanowicz, B.A., 2014. Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography, *Geophys. J. Int.,* **199**(3), 1303–1327.

Grand, S.P. & Helmberger, D.V., 1984. Upper mantle shear structure of North America, *Geophys. J. Int.,* **76**(2), 399–438.

Griffin, W., O'Reilly, S.Y., Doyle, B., Pearson, N., Coopersmith, H., Kivi, K., Malkovets, V. & Pokhilenko, N., 2004. Lithosphere mapping beneath the North American plate, *Lithos,* **77**(1), 873–922.

Gung, Y., Panning, M. & Romanowicz, B., 2003. Global anisotropy and the thickness of continents, *Nature,* **422**(6933), 707–711.

Heaman, L. & Kjarsgaard, B., 2000. Timing of eastern north american kimberlite magmatism: continental extension of the great meteor hotspot track?, *Earth planet. Sci. Lett.,* **178**(3), 253–268.

Heaman, L., Kjarsgaard, B. & Creaser, R., 2003. The timing of kimberlite magmatism in North America: implications for global kimberlite genesis and diamond exploration, *Lithos,* **71**(2), 153–184.

Herrin, E. & Taggart, J., 1968. Regional variations in P travel times, *Bull. seism. Soc. Am.,* **58**(4), 1325–1337.

Hoffman, P.F., 1988. United plates of America, the birth of a craton: early Proterozoic assembly and growth of laurentia, *Annu. Rev. Earth Planet. Sci.,* **16**(1), 543–603.

James, E.K., Dalton, C.A. & Gaherty, J.B., 2014. Rayleigh wave phase velocities in the Atlantic upper mantle, *Geochem. Geophys. Geosyst.,* **15**(11), 4305–4324.

Jones, A.G., Ledo, J., Ferguson, I.J., Craven, J.A., Unsworth, M.J., Chouteau, M. & Spratt, J.E., 2014. The electrical resistivity of Canada's lithosphere and correlation with other parameters: contributions from Lithoprobe and other programmes 1, *Can. J. Earth Sci.,* **51**(6), 573–617.

Jordan, T.H. & Paulson, E.M., 2013. Convergence depths of tectonic regions from an ensemble of global tomographic models, *J. geophys. Res.,* **118**(8), 4196–4225.

Kaban, M.K., Mooney, W.D. & Petrunin, A.G., 2015. Cratonic root beneath North America shifted by basal drag from the convecting mantle, *Nature Geosci.,* **8**(10), 797–800.

King, 2005. Archean cratons and mantle dynamics, *Earth planet. Sci. Lett.,* **234**(1–2), 1–14.

Komatitsch, D., 2002. The spectral-element method, beowulf computing, and global seismology, *Science,* **298**(5599), 1737–1742.

Kustowski, B., Ekström, G. & Dziewoński, A.M., 2008. Anisotropic shear-wave velocity structure of the earth's mantle: a global model, *J. geophys. Res.,* **113**(B6), doi:10.1029/2007JB005169.

Lee, C.-T.A., Luffi, P. & Chin, E.J., 2011. Building and destroying continental mantle, *Annu. Rev. Earth Planet. Sci.,* **39**(1), 59–90.

Lekić, V. & Romanowicz, B., 2011. Inferring upper-mantle structure by full waveform tomography with the spectral element method, *Geophys. J. Int.,* **185**(2), 799–831.

Lekić, V. & Romanowicz, B., 2011. Tectonic regionalization without a priori information: a cluster analysis of upper mantle tomography, *Earth planet. Sci. Lett.,* **308**(1), 151–160.

Levin, V., Menke, W. & Park, J., 1999. Shear wave splitting in the appalachians and the urals—a case for multilayered anisotropy, *J. geophys. Res.,* **104**(17), 975–17.

Li, X.-D. & Romanowicz, B., 1995. Comparison of global waveform inversions with and without considering cross-branch modal coupling, *Geophys. J. Int.,* **121**(3), 695–709.

Li, X.-D. & Romanowicz, B., 1996. Global mantle shear velocity model developed using nonlinear asymptotic coupling theory, *J. geophys. Res.,* **101**(B10), 22 245–22 272.

Love, A.E.H., 1911. *A Treatise on the Mathematical Theory of Elasticity,* Cambridge Univ. Press.

Marone, F. & Romanowicz, B., 2007. The depth distribution of azimuthal anisotropy in the continental upper mantle, *Nature,* **447**(7141), 198–201.

Marone, F., Gung, Y. & Romanowicz, B., 2007. Three-dimensional radial anisotropic structure of the north american upper mantle from inversion of surface waveform data, *Geophys. J. Int.,* **171**(1), 206–222.

Masson, Y. & Romanowicz, B., 2017a. Box tomography: localised imaging of remote targets buried in an unknown medium, a step forward for understanding key structures in the deep Earth, *Geophys. J. Int.*, doi:10.1093/gji/ggx141.

Masson, Y. & Romanowicz, B., 2017b. Fast computation of synthetic seismograms within a medium containing remote localized perturbations: a numerical solution to the scattering problem, *Geophys. J. Int.,* **208**(2), 674–692.

Masson, Y., Cupillard, P., Capdeville, Y. & Romanowicz, B., 2014. On the numerical implementation of time-reversal mirrors for tomographic imaging, *Geophys. J. Int.,* **196**(3), 1580–1599.

Mégnin, C. & Romanowicz, B., 1999. The effects of the theoretical formalism and data selection on mantle models derived from waveform tomography, *Geophys. J. Int.,* **138**(2), 366–380.

Mégnin, C. & Romanowicz, B., 2000. The three-dimensional shear velocity structure of the mantle from the inversion of body, surface and higher-mode waveforms, *Geophys. J. Int.,* **143**(3), 709–728.

Mochizuki, E., 1986. The free oscillations of an anisotropic and heterogeneous earth, *Geophys. J. Int.,* **86**(1), 167–176.
Mocquet, A. & Romanowicz, B., 1990. Three-dimensional structure of the upper mantle beneath the Atlantic ocean inferred from long period Rayleigh waves, ii, inversion, *J. geophys. Res,* **95,** 6787–6798.
Montagner, J.-P. & Anderson, D.L., 1989. Petrological constraints on seismic anisotropy, *Phys. Earth planet. Inter.,* **54**(1–2), 82–105.
Monteiller, V., Chevrot, S., Komatitsch, D. & Wang, Y., 2015. Three-dimensional full waveform inversion of short-period teleseismic wavefields based upon the SEM–DSM hybrid method, *Geophys. J. Int.,* **202**(2), 811–827.
Németh, B., Clowes, R.M. & Hajnal, Z., 2005. Lithospheric structure of the Trans-Hudson Orogen from seismic refraction-wide-angle reflection studies, *Can. J. Earth Sci.,* **42**(4), 435–456.
Nettles, M. & Dziewoński, A.M., 2008. Radially anisotropic shear velocity structure of the upper mantle globally and beneath North America, *J. geophys. Res.,* **113**(B2), doi:10.1029/2006JB004819.
Panning, M. & Romanowicz, B., 2006. A three-dimensional radially anisotropic model of shear velocity in the whole mantle, *Geophys. J. Int.,* **167**(1), 361–379.
Porritt, R.W., Allen, R.M. & Pollitz, F.F., 2014. Seismic imaging east of the Rocky Mountains with USarray, *Earth planet. Sci. Lett.,* **402,** 16–25.
Porritt, R.W., Miller, M.S. & Darbyshire, F.A., 2015. Lithospheric architecture beneath Hudson Bay, *Geochem. Geophys. Geosyst.,* **16**(7), 2262–2275.
Poupinet, G., 1979. On the relation between P-wave travel time residuals and the age of continental plates, *Earth planet. Sci. Lett.,* **43**(1), 149–161.
Pratt, R.G., Shin, C. & Hick, G., 1998. Gauss–Newton and full newton methods in frequency–space seismic waveform inversion, *Geophys. J. Int.,* **133**(2), 341–362.
Rader, E., Emry, E., Schmerr, N., Frost, D., Cheng, C., Menard, J., Yu, C.-Q. & Geist, D., 2015. Characterization and petrological constraints of the midlithospheric discontinuity, *Geochem. Geophys. Geosyst.,* **16**(10), 3484–3504.
Ritsema, J., Deuss, A., van Heijst, H.J. & Woodhouse, J.H., 2010. S40rts: a degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements, *Geophys. J. Int.,* **184**(3), 1223–1236.
Romanowicz, B., 1987. Multiplet-multiplet coupling due to lateral heterogeneity: asymptotic effects on the amplitude and frequency of the earth's normal modes, *Geophys. J. Int.,* **90**(1), 75–100.
Romanowicz, B.A., 1979. Seismic structure of the upper mantle beneath the United States by three-dimensional inversion of body wave arrival times, *Geophys. J. Int.,* **57**(2), 479–506.
Romanowicz, B.A., Panning, M.P., Gung, Y. & Capdeville, Y., 2008. On the computation of long period seismograms in a 3-D earth using normal mode based approximations, *Geophys. J. Int.,* **175**(2), 520–536.
Rondenay, S., Bostock, M.G., Hearn, T.M., White, D.J. & Ellis, R.M., 2000. Lithospheric assembly and modification of the SE canadian shield: Abitibi-Grenville teleseismic experiment, *J. geophys. Res.,* **105**(B6), 13735–13754.
Schaeffer, A. & Lebedev, S., 2014. Imaging the north american continent using waveform inversion of global and USArray data, *Earth planet. Sci. Lett.,* **402,** 26–41.
Schaeffer, A.J. & Lebedev, S., 2013. Global shear speed structure of the upper mantle and transition zone, *Geophys. J. Int.,* **194**(1), 417–449.
Shapiro, N.M. & Ritzwoller, M.H., 2002. Monte-Carlo inversion for a global shear-velocity model of the crust and upper mantle, *Geophys. J. Int.,* **151**(1), 88–105.
Sigloch, K., 2011. Mantle provinces under North America from multifrequency P wave tomography, *Geochem. Geophys. Geosyst.,* **12**(2), doi:10.1029/2010GC003421.
Silveira, G. & Stutzmann, E., 2002. Anisotropic tomography of the Atlantic ocean, *Phys. Earth planet. Inter.,* **132**(4), 237–248.
Snyder, D.B., Hillier, M.J., Kjarsgaard, B.A., de Kemp, E.A. & Craven, J.A., 2014. Lithospheric architecture of the slave craton, northwest Canada, as determined from an interdisciplinary 3-D model, *Geochem. Geophys. Geosyst.,* **15**(5), 1895–1910.
Su, W.-j., Woodward, R.L. & Dziewonski, A.M., 1994. Degree 12 model of shear velocity heterogeneity in the mantle, *J. geophys. Res.,* **99**(B4), 6945–6980.
Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2010. Seismic tomography of the Southern California crust based on spectral-element and adjoint methods, *Geophys. J. Int.,* **180**(1), 433–462.
Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation,* SIAM.
Tarantola, A. & Valette, B., 1982. Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys.,* **20**(2), 219–232.
Thybo, H. & Perchuć, E., 1997. The seismic 8 discontinuity and partial melting in continental mantle, *Science,* **275**(5306), 1626–1629.
Unser, M., Aldroubi, A. & Eden, M., 1993a. B-spline signal processing. I. Theory, *IEEE Trans. Signal Process.,* **41**(2), 821–833.
Unser, M., Aldroubi, A. & Eden, M., 1993b. B-spline signal processing. II. Efficiency design and applications, *IEEE Trans. Signal Process.,* **41**(2), 834–848.
van der Lee, S. & Frederiksen, A., 2005. Surface wave tomography applied to the North American upper mantle, in *Seismic Earth: Array Analysis of Broadband Seismograms,* pp. 67–80, eds Alan, L. & Guust, N., Wiley-Blackwell.
van der Lee, S. & Nolet, G., 1997. Upper mantle S velocity structure of North America, *J. geophys. Res.,* **102**(B10), 22 815–22 838.
Villemaire, M., Darbyshire, F. & Bastow, I., 2012. P-wave tomography of eastern North America: evidence for mantle evolution from Archean to Phanerozoic, and modification during subsequent hot spot tectonism, *J. geophys. Res.,* **117**(B12), doi:10.1029/2012JB009639.
Wang, Y. *et al.*, 2016. The deep roots of the western pyrenees revealed by full waveform inversion of teleseismic p waves, *Geology,* **44**(6), 475–478.
Wang, Z. & Dahlen, F.A., 1995. Spherical-spline parameterization of three-dimensional earth models, *Geophys. Res. Lett.,* **22**(22), 3099–3102.
Whitmeyer, S.J. & Karlstrom, K.E., 2007. Tectonic model for the Proterozoic growth of North America, *Geosphere,* **3**(4), 220–259.
Woodhouse, J.H. & Dziewonski, A.M., 1984. Mapping the upper mantle: three-dimensional modeling of earth structure by inversion of seismic waveforms, *J. geophys. Res.,* **89**(B7), 5953–5986.
Yuan, H. & Romanowicz, B., 2010. Lithospheric layering in the North American craton, *Nature,* **466**(7310), 1063–1068.
Yuan, H., Romanowicz, B., Fischer, K.M. & Abt, D., 2011. 3-D shear wave radially and azimuthally anisotropic velocity model of the north american upper mantle, *Geophys. J. Int.,* **184**(3), 1237–1260.
Yuan, H., French, S., Cupillard, P. & Romanowicz, B., 2014. Lithospheric expression of geological units in central and eastern North America from full waveform tomography, *Earth planet. Sci. Lett.,* **402,** 176–186.
Zhou, Y., Dahlen, F., Nolet, G. & Laske, G., 2005. Finite-frequency effects in global surface-wave tomography, *Geophys. J. Int.,* **163**(3), 1087–1111.
Zhu, H., Bozdağ, E. & Tromp, J., 2015. Seismic structure of the european upper mantle based on adjoint tomography, *Geophys. J. Int.,* **201**(1), 18–52.
Zhu, H., Komatitsch, D. & Tromp, J., 2017. Radial anisotropy of the north american upper mantle based on adjoint tomography with USarray, *Geophys. J. Int.,* **211**(1), 349–377.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

APPENDIX: RESOLUTION TESTS

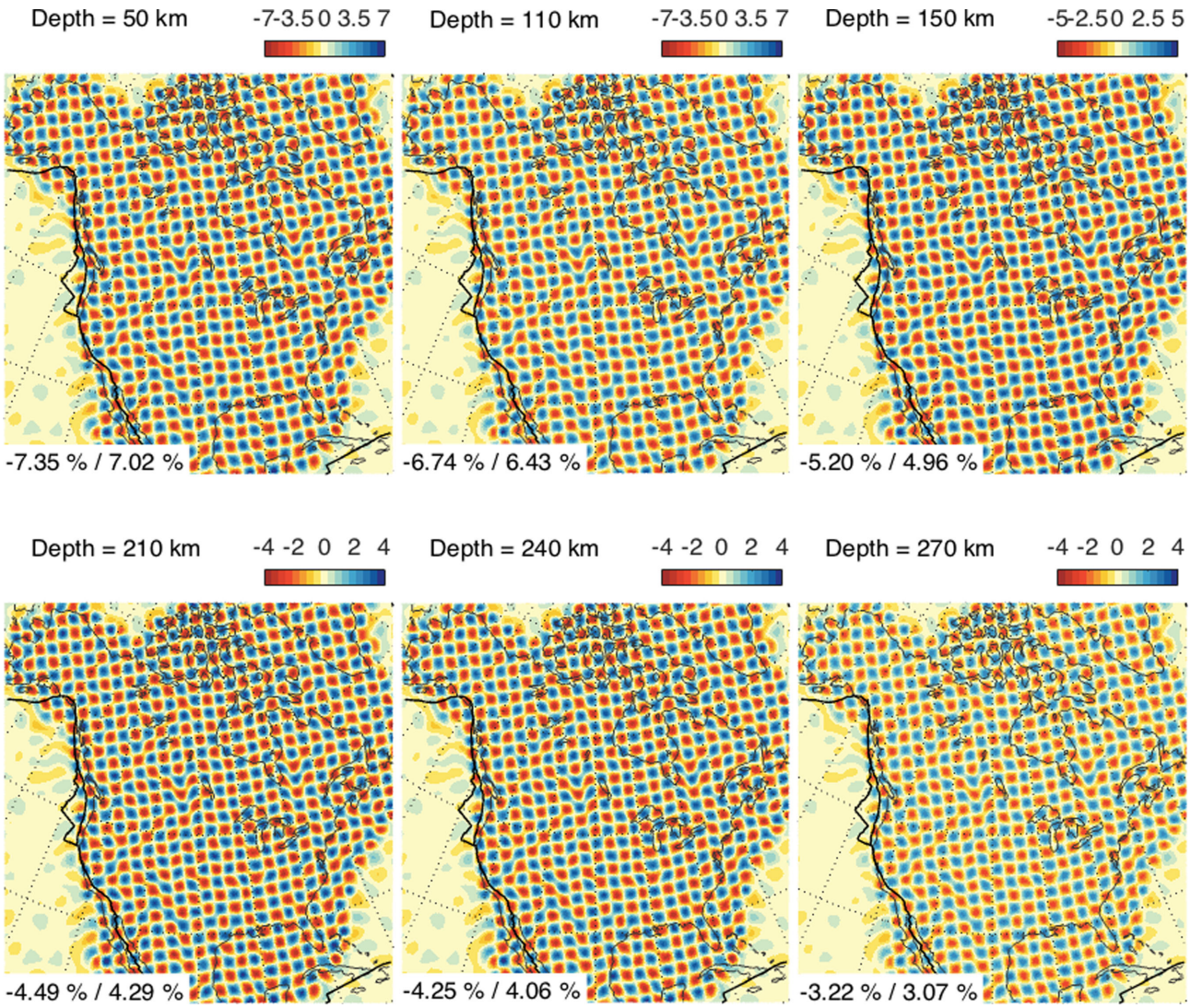

Figure A1. Input model for checkerboard tests of V_S in North America. Each checkerboard test model is constructed by assigning a non-zero coefficient to a single radial spline $v_q(r)$ (eq. 6), resulting in 3-D structure concentrated around the corresponding depth (shown at the top left). The lateral structure is obtained by assigning the same value but opposite signs to each spherical spline $\beta_p(\theta, \varphi)$ (eq. 6) with 2° node spacing. For each model, the structure is shown only at the depth corresponding to the maximum velocity perturbation. The colour bars, at the top right of each panel, indicates perturbations in per cent. Note that the scale is different in different panels. Minimum and maximum perturbations are indicated at the bottom of each map. Amplitudes of cells mimic amplitudes observed in the model presented in this study.

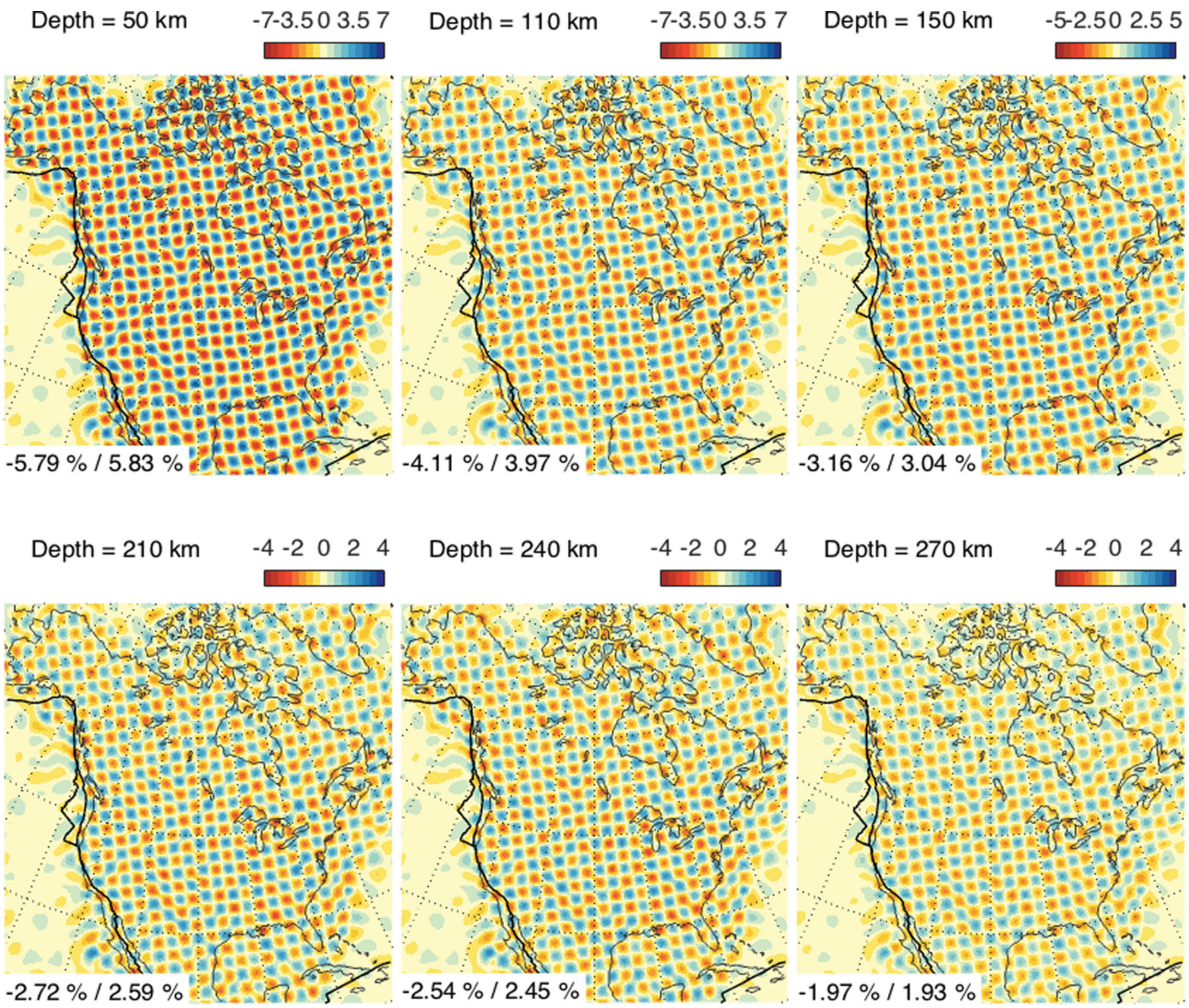

Figure A2. Output model for checkerboard tests of V_S in North America, corresponding to the input models shown in Fig A1. Presentation is as in Fig. A1. We calculate the resolving power of the inversion by using classical linear resolution analysis (e.g. Tarantola 2005), whereby we construct the resolution matrix corresponding to the last iteration of our inversion. While structures show good spatial recovery, amplitudes – due to application of damping during the inversion – is generally reduced. The recovery is poorer in the northern part of the continent (between Greenland and North America especially) as expected from our density and azimuthal coverage (see Fig. 4). In the future, the use of out-of-box stations in our Box Tomography inversion procedure will greatly improve the illumination of this area while keeping its cost low.

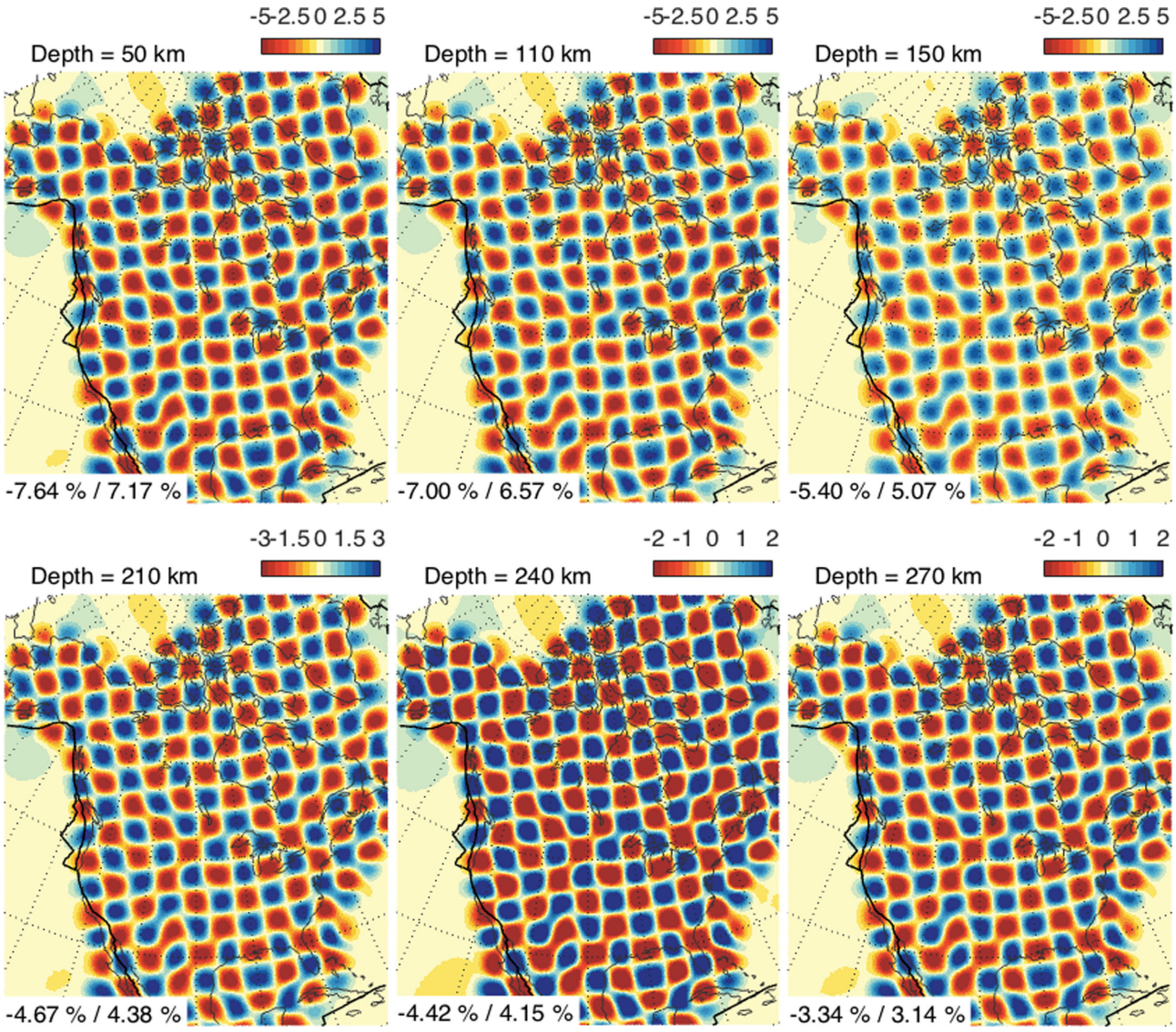

Figure A3. Input model for checkerboard tests of ξ in North America. Presentation is as in Fig. A1. The input models are constructed in the same way as for V_S, but the spherical spline node spacing is here 4°.

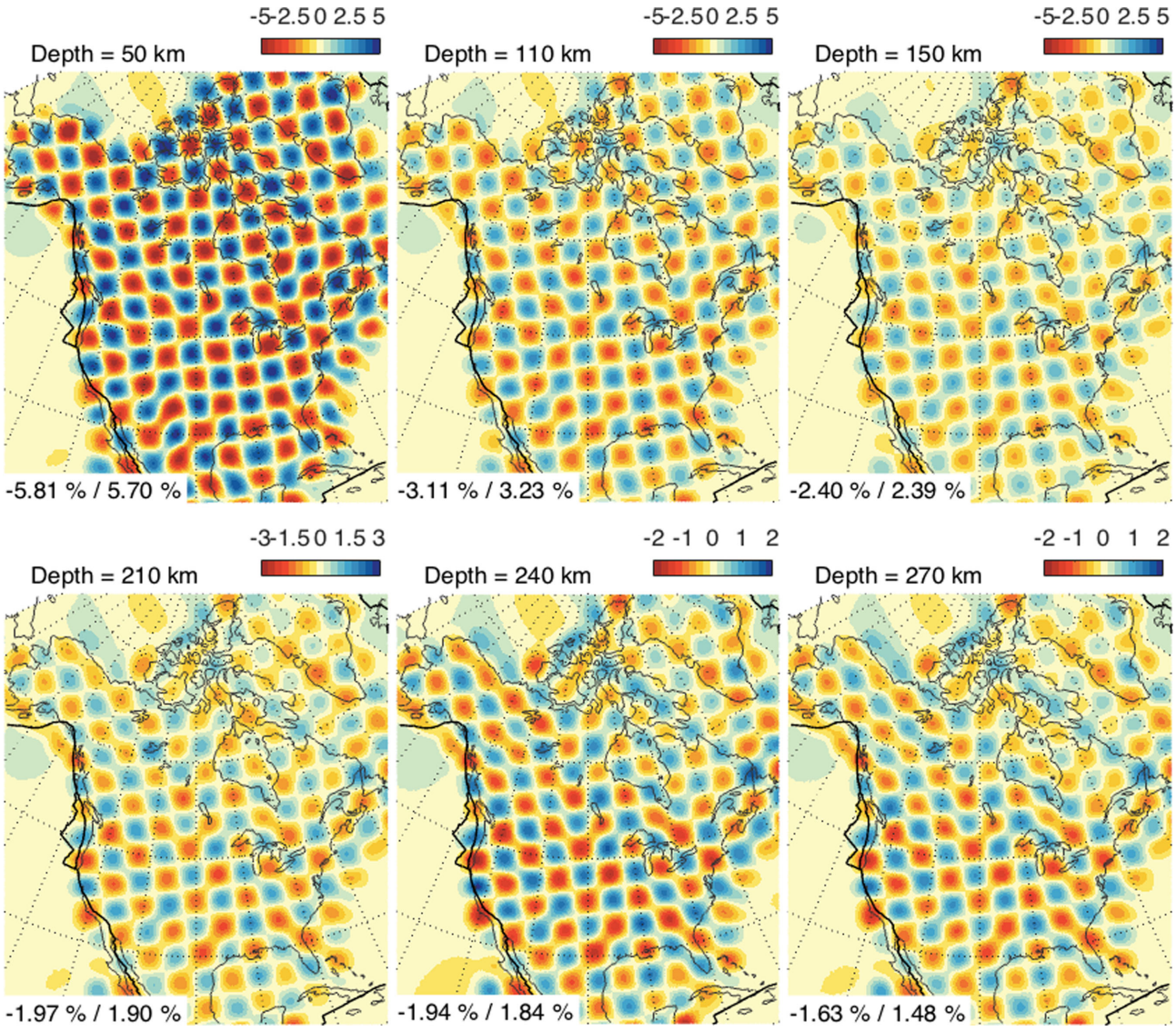

Figure A4. Output model from checkerboard tests of ξ in North America. Same as in Fig. A2.

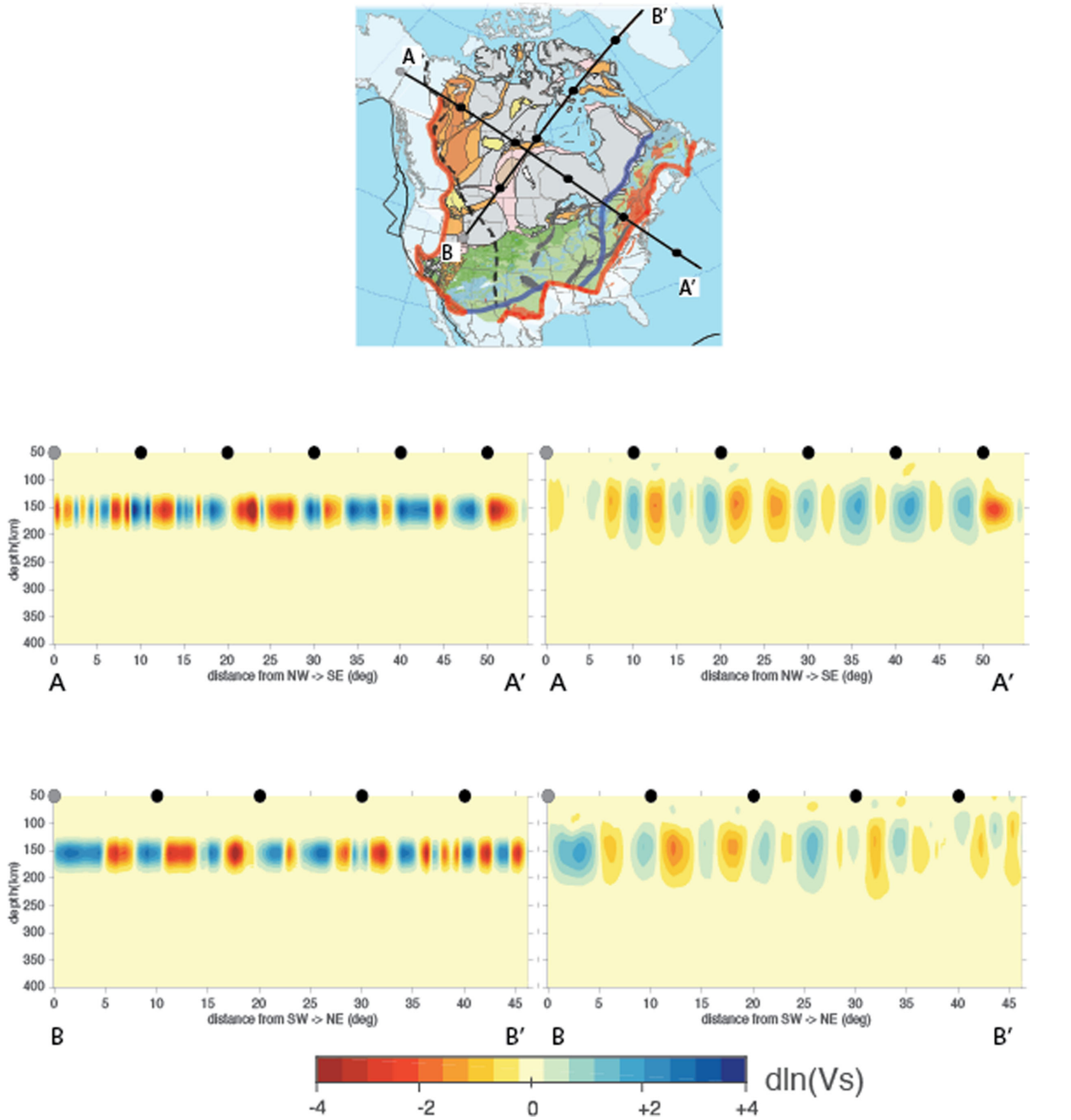

Figure A5. Depth cross-sections from 50 to 400 km depth for checkerboard tests of V_S centred at 150 km depth in North America. Perturbations of V_S are shown according to the colour bar at the bottom of the figure. The left-hand panels show the input model and the right-hand panels show the output model. The input and output models are constructed as in Figs A1 and A2. The location of the different depth profiles shown is indicated on the geological map at the top. Small black circles at top of each cross-section correspond to the same black circles on the geological map.

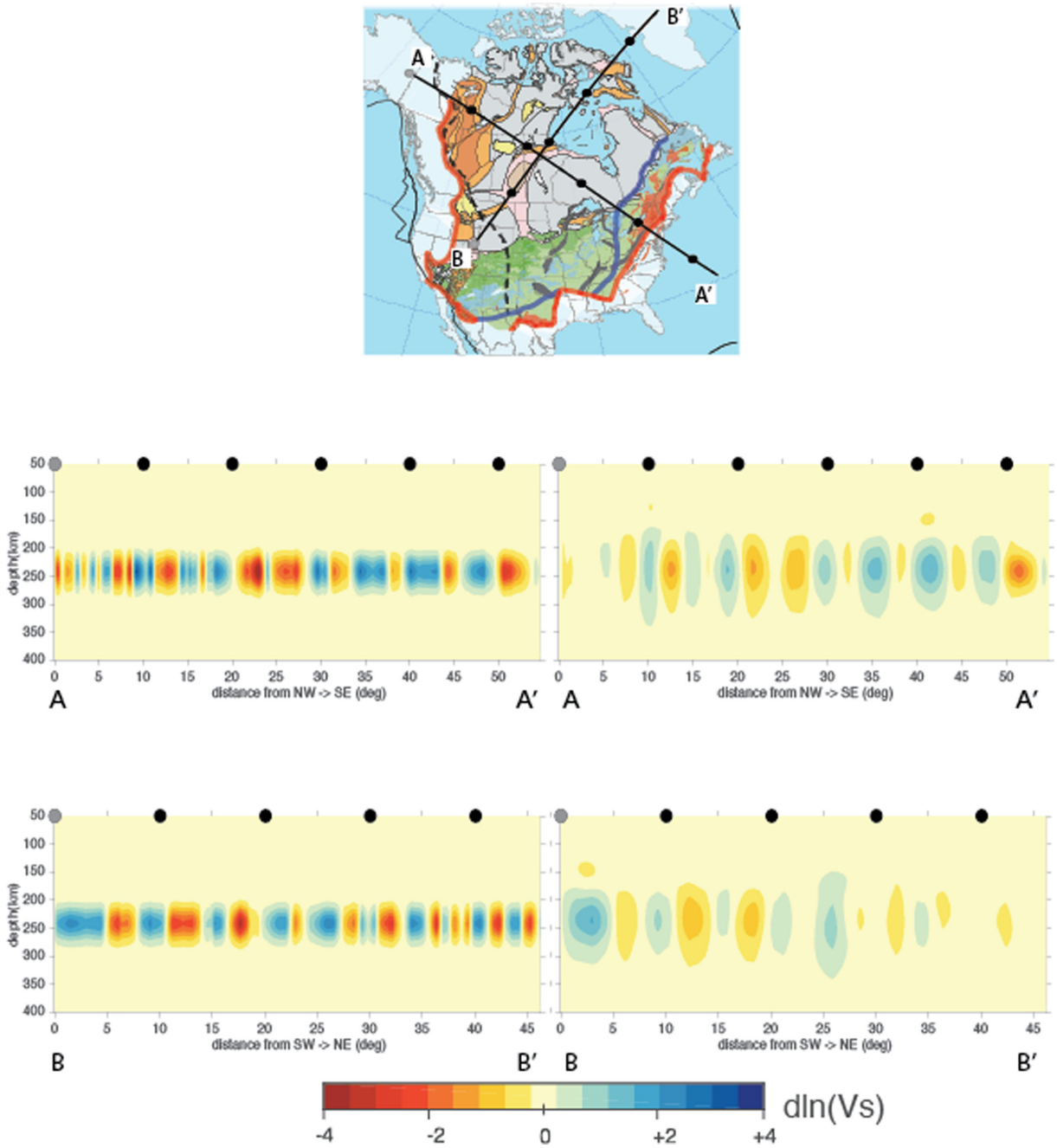

Figure A6. Depth cross-sections from 50 to 400 km depth for the output of the checkerboard test for V_S centred at 240 km depth in North America. Same as in Fig. A5.

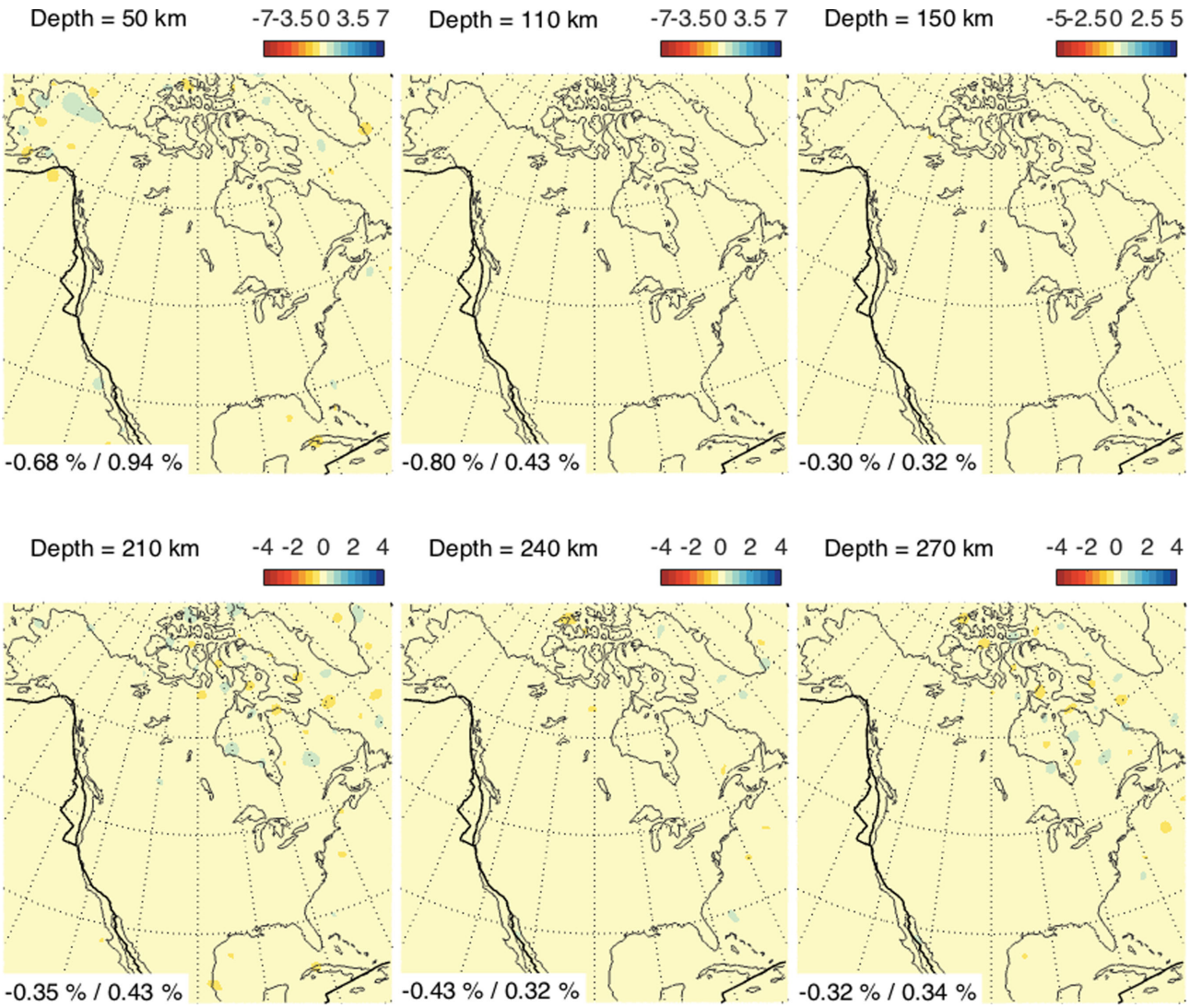

Figure A7. Results of checkerboard test for cross-contamination between ξ and V_s. Here, the input model has structure in ξ, as shown in Fig A3, and heterogeneity in V_S is set to zero. The presentation is as in Fig A2. It is clear that ξ does not contaminate V_S except for some very small anomalies $d\ln(V_S) < 1$ per cent in the northern part of the inverted region, where the recovery from the V_S only resolution test is poor (see Fig. A2) and path density and azimuthal coverage is also poor (see Fig. 4).

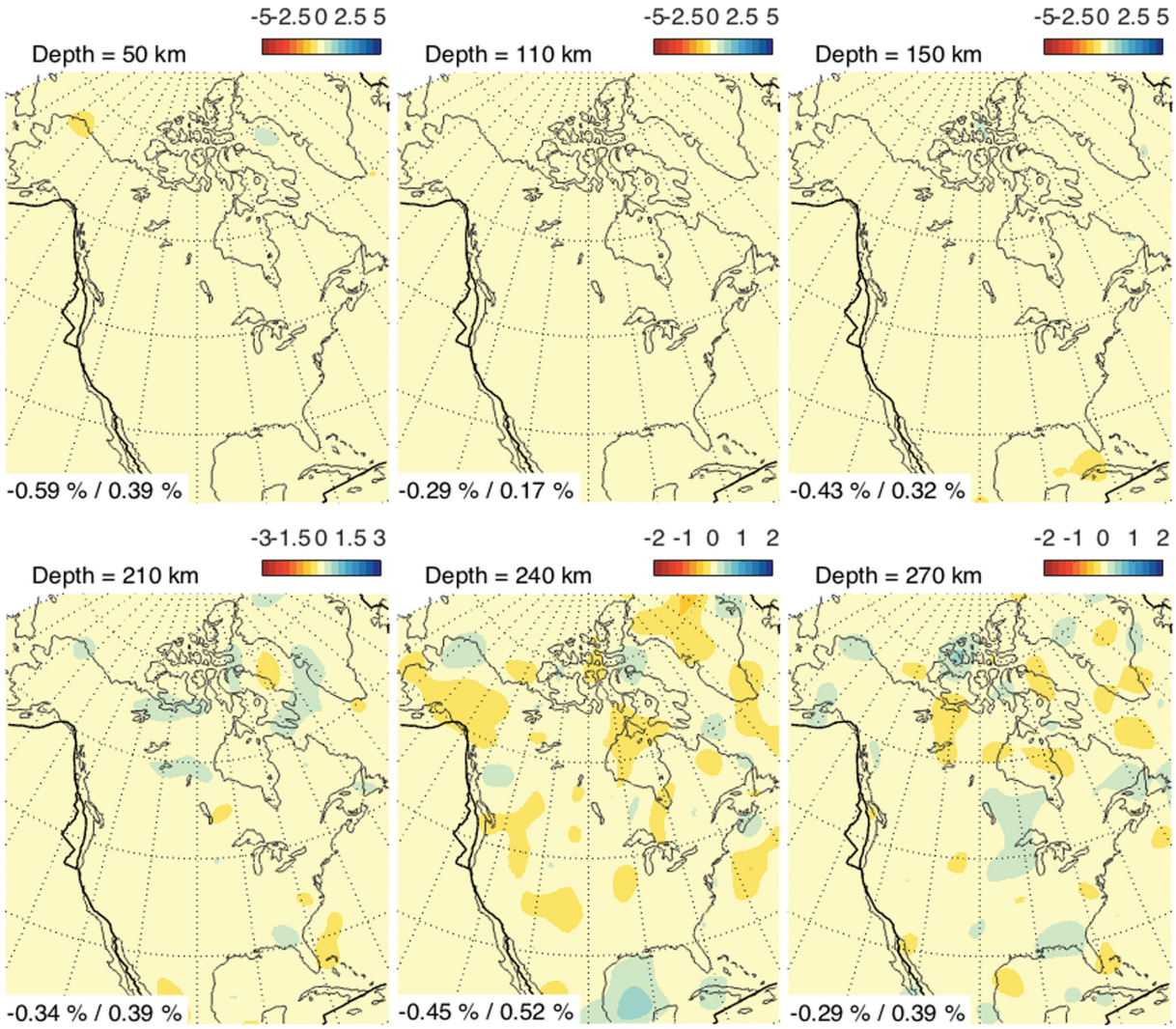

Figure A8. Same as Fig. A7 except that the input model now has structure in V_S (input model is shown in Fig. A1) and heterogeneity in ξ is set to zero. Presentation is as in Fig. A4. Unlike V_S, ξ structure is somewhat contaminated by V_S at depths greater than 200 km. Still, anomalies due to contamination are small $d\ln(\xi) < 0.5$ per cent and correspond to depths where the amplitude recovery from the resolution in Fig. A4 is below 50 per cent.